NUSC Technical Report 4243

Operating Characteristics for Maximum Likelihood Detection of Signals in Gaussian Noise Of Unknown Level
I. Coherent Signals of Unknown Level

Albert H. Nuttall
Peter G. Cable
Office of the Director of Science and Technology

27 March 1972

NAVAL UNDERWATER SYSTEMS CENTER

Approved for public release; distribution unlimited.

ADMINISTRATIVE INFORMATION

This report was prepared under NUSC Project No. A-752-05, Navy Subproject No. ZFXX212001, "Statistical Communication with Applications to Sonar Signal Processing," Principal Investigator, Dr. A. H. Nuttall (Code TC). The sponsoring activity is Chief of Naval Material, Program Manager, Dr. J. H. Huth.

The Technical Reviewer for this report was Dr. D. W. Hyde, Code TC.

REVIEWED AND APPROVED: 27 March 1972

W.A. Von Winkle

W. A. Von Winkle
Director of Science and Technology

Inquiries concerning this report may be addressed to the authors,
New London Laboratory, Naval Underwater Systems Center,
New London, Connecticut 06320

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Underwater Systems Center Newport, Rhode Island 02840	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE
OPERATING CHARACTERISTICS FOR MAXIMUM LIKELIHOOD DETECTION OF SIGNALS IN GAUSSIAN NOISE OF UNKNOWN LEVEL
I. COHERENT SIGNALS OF UNKNOWN LEVEL

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Research Report

5. AUTHOR(S) (First name, middle initial, last name)
Albert H. Nuttall
Peter G. Cable

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
27 March 1972	68	14

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. A-752-05	TR 4243
c. ZFXX212 001	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Department of the Navy

13. ABSTRACT

The detection and false-alarm probabilities for maximum likelihood detection of coherent signals in noise, where the signal and noise levels are unknown, are derived and evaluated. The maximum likelihood detector is shown to be optimum in the following sense: out of the class of processors that yields a specified false-alarm probability without knowledge of the noise level, it is uniformly most powerful for unknown (positive) signal level.

For M samples of signal-plus-noise or noise, and N samples of noise-alone, curves of detection probability versus signal-to-noise ratio, with false-alarm probability as a parameter, are presented for values of M + N ranging from 2 to 27. Curves of signal-to-noise ratio required for 0.5 detection probability are presented for several values of false-alarm probability.

Detection performance improves very rapidly as M + N increases from 2 to 4 or 5; the degree of improvement for larger M + N depends on the exact value of false-alarm probability desired. For example, small false-alarm probabilities such as 10^{-6} require ten or more total samples before the performance improvement begins to slacken with the number of samples taken.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Operating Characteristics						
Maximum Likelihood Detection						
Unknown Noise Level						
Uniformly Most Powerful						
Scaling Invariance						
Parametric Detection						
Constant False Alarm Receiver						
Normalization						

ABSTRACT

The detection and false-alarm probabilities for maximum likelihood detection of coherent signals in noise, where the signal and noise levels are unknown, are derived and evaluated. The maximum likelihood detector is shown to be optimum in the following sense: out of the class of processors that yields a specified false-alarm probability without knowledge of the noise level, it is uniformly most powerful for unknown (positive) signal level.

For M samples of signal-plus-noise or noise, and N samples of noise-alone, curves of detection probability versus signal-to-noise ratio, with false-alarm probability as a parameter, are presented for values of M + N ranging from 2 to 27. Curves of signal-to-noise ratio required for 0.5 detection probability are presented for several values of false-alarm probability.

Detection performance improves very rapidly as M + N increases from 2 to 4 or 5; the degree of improvement for larger M + N depends on the exact value of false-alarm probability desired. For example, small false-alarm probabilities such as 10^{-6} require ten or more total samples before the performance improvement begins to slacken with the number of samples taken.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ABBREVIATIONS

CF	Characteristic Function
GLR	Generalized Likelihood Ratio
LR	Likelihood Ratio
ML	Maximum Likelihood
PDF	Probability Density Function
rms	Root-mean-square
RV	Random Variable
SNR	Signal-to-Noise Ratio
UMP	Uniformly Most Powerful

LIST OF SYMBOLS

d_T	Signal-to-noise ratio (voltage), Eqs. (15) - (17)
H_o, H_1	Signal absent and signal present hypotheses
K	M+N-2
m	Mean, signal strength
m_1	Sample mean, maximum likelihood estimate
M	Number of potential signal-plus-noise samples
n	Integer
N	Number of noise-alone samples
P_D	Probability of detection
P_F	Probability of false alarm
p_o, p_1	Probability density functions under H_o, H_1
r	Scale factor in test, Eq. (10)
R	Observation vector
x_i	i-th sample of signal-plus-noise
y_j	j-th sample of noise-alone
$\alpha_{1,2}$	Constants
σ^2	Noise variance
σ_o^2, σ_1^2	Sample variances, maximum likelihood estimates
Φ	Cumulative Gaussian distribution, Eq. (5)
Λ, T, $\nu_{1,2}$, γ	Thresholds

OPERATING CHARACTERISTICS FOR MAXIMUM LIKELIHOOD DETECTION OF SIGNALS IN GAUSSIAN NOISE OF UNKNOWN LEVEL

I. COHERENT SIGNALS OF UNKNOWN LEVEL

INTRODUCTION

Signal-detection systems often are required to operate in environments where the background noise level is subject to unknown variations. For example, receivers for active sonar (or radar) are called upon to process returns against interference that includes time-varying reverberation (or clutter). From a practical point of view, it can be expected that both the level and the spectrum of the interference will be unknown and changing with time. In such a situation, establishing a noise-level reference for each range (time delay) increment is obviously necessary if a specified probability of false alarm P_F is to be maintained. The subject of this report is a study of a detector that maximizes the probability of detection P_D and at the same time maintains a desired P_F in the presence of Gaussian noise of unknown level and spectral density (or autocovariance function).

Signal-detection applications in a stationary environment of unknown noise power level present a potentially simpler problem than the one addressed here. In the stationary environment it is likely that the unknown noise parameters can be estimated independently of the signal detection procedure. Then, prior to detection, a noise-reference level could be established to any degree of accuracy, and the desired false-alarm probability could be approximately realized.

For the more realistic situation of signal detection in a nonstationary environment with noise of unknown spectral density, a fixed detection-threshold would yield a varying false-alarm probability dependent on changes in the noise level and spectral density. This undesirable feature can be alleviated (at the expense of introducing additional noise into the detection procedure) by estimating the noise level at each increment of time in order to establish a detection threshold that tends to track the noise power level. It is often the case that the noise background is varying slowly enough in time that independent samples of noise-alone, from the same ensemble as the sample being tested for signal presence, are available in neighboring time increments. Obviously, extra samples of noise-alone can be used to improve the estimate of noise level, and allowance for them is included in this investigation.

It is worthwhile remarking that this report is applicable to problems of changing noise environments other than just temporally nonstationary examples. For example, passive receiving arrays generally operate in spatially nonisotropic noise fields; if the samples of signal-plus-noise and noise-alone are assumed to be obtained from neighboring-look directions, the results of this investigation are applicable in a spatial normalization context. Or, if a system for detecting narrowband components in a broadband input is considered, the samples may be assumed to come from adjacent frequency bins, and the results of this investigation are also applicable in the frequency domain. To avoid an unnecessarily cumbersome discussion, however, the terminology of temporal nonstationarity will be employed throughout the remainder of the report.

It will be recognized that inability to specify the spectral density of the noise places analysis within the framework of a nonparametric detection problem, for which the goal will be to achieve a constant false-alarm-rate receiver. In the following discussion, it will be assumed that the input to the receiving system has been filtered to remove noise outside the signal band and that the prefiltered input is sampled at times sufficiently separated for the samples to be statistically independent. The approach will also assume that the noise has a specific statistical structure, namely, Gaussian, thereby casting the originally nonparametric problem into a parametric form (although the parameters appearing in the probability density functions (PDFs) are unknown).

In the general case, the values of the unknown parameters can be estimated by separately maximizing the PDFs for the signal present hypothesis H_1 and signal absent hypothesis H_0. The likelihood ratio (LR) formed under these conditions (principle of maximum likelihood) can be used to obtain a generalized likelihood ratio (GLR) test. This procedure was carried through by Helstrom[1] for coherent signals in Gaussian noise of unknown level and spectral density (no extra noise samples were assumed present in his analysis) and resulted in a t-statistic test. In the coherent detection case, however, a more powerful result than Helstrom's can be obtained, as shown by Scharf and Lytle[2] and as derived in Appendix A to this report. The resultant t-test is, with unknown signal level, uniformly most powerful (UMP) of all tests which maintain a specified P_F (irrespective of noise level changes) and is optimum (in a Neyman-Pearson sense) for the same (scaling invariance) class of tests when the signal level is known.

The problem of detecting a signal of known form (but possibly unknown level) in Gaussian noise of unknown level has been the subject of a number of investigations. In a parametric context, Helstrom[1] has considered the question of detecting signals in noise of unknown level but known normalized autocovari-

ance function. In a series of problems, he employs a method of solution based on a maximum likelihood estimate of the noise level, to derive a t-statistic test whose asymptotic reliability is as good as that of the known-noise receiver; physical realization of this processor is not practical, however. Scharf and Lytle have shown that the t-test is UMP-invariant. In the nonparametric situation (unknown noise autocovariance), Carlyle and Thomas[3] have treated the t-statistic as an asymptotically nonparametric detector. Other authors have considered receivers that employ a separate (extra) input exhibiting the same kind of noise as that which corrupts the signal but is known to contain no signal. Capon,[4] Hancock and Lainiotis,[5] and Carlyle and Thomas[3] have analyzed a number of nonparametric receivers adapted from so-called Wilcoxon tests (both with and without extra noise-alone samples) and have obtained asymptotic relative efficiencies for several such detectors. Finally, Spooner,[6] using an averaged likelihood ratio test, has treated the problem of detecting signals in white Gaussian noise whose unknown random level is assumed to be gamma distributed.

This report presents a self-contained statement, including descriptions and derivations, about the maximum likelihood and optimum scaling-invariant procedures referred to above. The analyses in the following sections not only constitute a review of previous work on coherent detection of signals in noise of unknown level, but they also provide a foundation for consideration of phase-incoherent detection of deterministic signals and for consideration of detection of stochastic signals in noise of unknown level. In two important respects, however, the present effort represents extensions of previous work. First, allowance is made for including extra samples of noise-alone in the maximum likelihood estimates of the noise power. Second, extensive numerical results are given for the performance of the maximum likelihood (optimum) detector, both with and without extra noise-alone samples.

PROBLEM STATEMENT

We have available M samples of either signal-plus-noise or noise, and N samples of noise-alone, denoted by $\{x_i\}_1^M$ and $\{y_j\}_1^N$, respectively. The observation "vector" R is made up of these M + N samples:

$$R^T \equiv [x_1 \cdots x_M \, y_1 \cdots y_N] . \tag{1}$$

The samples are assumed to be statistically independent of each other, and the noise is Gaussian. The PDF of R under H_0 is

$$p_o(R) = \prod_{i=1}^{M} \left\{(2\pi\sigma^2)^{-1/2} \exp\left(-\frac{x_i^2}{2\sigma^2}\right)\right\} \prod_{j=1}^{N} \left\{(2\pi\sigma^2)^{-1/2} \exp\left(-\frac{y_j^2}{2\sigma^2}\right)\right\}, \qquad (2)$$

where it is assumed that the noise variance σ^2 is the same for all M+N samples. The PDF of R under H_1 is

$$p_1(R) = \prod_{i=1}^{M} \left\{(2\pi\sigma^2)^{-1/2} \exp\left(-\frac{(x_i-m)^2}{2\sigma^2}\right)\right\} \prod_{j=1}^{N} \left\{(2\pi\sigma^2)^{-1/2} \exp\left(-\frac{y_j^2}{2\sigma^2}\right)\right\}, \qquad (3)$$

where it is also assumed that the signal strength, as reflected by the mean m, is the same for all M signal-plus-noise samples.

Various possible cases of knowledge about m and σ^2 are summarized in Table 1. For cases (b) and (d), where m is unknown, it is assumed that m is known to be positive. This means that phase-coherent processing would be required in the narrowband signal-detection mode; a phase-tracking device would be needed in practice. The quantities $\alpha_{1,2}$ are constants, which are adjusted to realize a prescribed P_F.

The optimum processor for case (a), in the sense of maximum P_D for fixed P_F [Ref. 7, pp. 27-28], is indicated in the table. Its performance capability is governed by

$$P_F = \Phi(-\Lambda), \quad P_D = \Phi(\sqrt{M}\,\frac{m}{\sigma} - \Lambda), \qquad (4)$$

where Λ is a normalized threshold, and the cumulative Gaussian distribution

$$\Phi(x) \equiv \int_{-\infty}^{x} dt\,(2\pi)^{-1/2} \exp(-t^2/2). \qquad (5)$$

A desired P_F can be maintained by choice of Λ since σ is known. The processor for case (b) is identical to that for case (a), and is UMP [Ref. 7, pp. 88-92]; its performance is identidal to that given in (4). Samples $\{y_j\}_1^N$ are of no use in cases (a) and (b).

Table 1

FOUR CASES OF COHERENT PROCESSORS

Case	Mean m	Variance σ^2	Receiver Processing	
(a)	known	known	$\sum_{i=1}^{M} x_i \gtrless \alpha_1$	Optimum
(b)	unknown (positive)	known	$\sum_{i=1}^{M} x_i \gtrless \alpha_1$	UMP
(c)	known	unknown	$\sum_{i=1}^{M} x_i \gtrless \alpha_2 \left(\sum_{i=1}^{M} x_i^2 + \sum_{j=1}^{N} y_j^2 \right)^{1/2}$	Optimum under scaling invariance
(d)	unknown (positive)	unknown	$\sum_{i=1}^{M} x_i \gtrless \alpha_2 \left(\sum_{i=1}^{M} x_i^2 + \sum_{j=1}^{N} y_j^2 \right)^{1/2}$	UMP under scaling invariance
			Note: If the upper inequality is satisfied, declare H_1. If the lower inequality is satisfied, declare H_0.	

Cases (c) and (d) have unknown noise variance σ^2 and are the situations of interest. Two approaches have been considered. The first is to employ the principle of maximum likelihood (ML) to estimate the unknown parameters and to use the GLR.[1,7] In case (c) the resulting test has the disadvantageous feature that the precise value of m must be known in order to realize prescribed P_F. In case (d) the resultant ML test is as indicated in Table 1. (Derivations for these two cases are presented in the next section.)

The second approach is based upon a scaling invariance developed in Appendix A [see also Ref. 8, ch. 6]. Basically, the idea is to consider the (infinite) class of processors that realize a specified P_F without knowledge of the noise level σ^2, and to select that processor which maximizes P_D. The resultant processors for cases (c) and (d) are identical and are given in Table 1; for case (c) the processor is optimum under the scaling invariance requirement, and for case (d) it is the UMP processor under scaling invariance. The processor is similar to that considered by Scharf and Lytle.[2]

Thus the ML-principle and the scaling-invariance approaches yield the same processor for case (d). The two approaches yield different processors for case (c). We have selected the processor indicated in Table 1 for both cases (c) and (d) because it realizes prescribed P_F without knowledge of m and σ^2 and is optimum in the sense of realizing maximum P_D. The noise-alone measurements $\{y_j\}_1^N$ enter through the sample power in the measurements.

Case (c) is somewhat unrealistic in that the signal strength is assumed known, while the noise level is not. Case (d) is an often encountered situation resulting, for example, from a nonstationary noise background and a channel with unknown or time-varying attenuation. Nevertheless, since the processors for both cases are identical under the scaling invariance requirement, the analysis and performance to be presented apply equally well to both cases. The actually attained P_D will depend upon the true values of m and σ^2; these, in turn, depend on the detailed signals transmitted, the noise power received, and the receiver filtering. The present analysis will indicate what value of m/σ is necessary in order to realize, for a wide range of M and N, prescribed P_F and P_D.

Since the processors for cases (c) and (d) are required to guarantee a specified P_F without knowledge of σ^2, they will not perform as well as those for cases (a) and (b), which have knowledge of σ^2. The additional signal-to-noise ratio (SNR) required will be part of the results of this investigation.

PROBLEM SOLUTION

For case (d), m and σ^2 in (2) and (3) are unknown; however, m is known to be nonnegative. According to the procedure for utilizing ML estimates, we choose σ^2 in (2) to maximize $p_0(\mathbf{R})$ for a given observation $\mathbf{R}$, and we choose m and σ^2 in (3) to maximize $p_1(\mathbf{R})$. Maximization of $p_0(\mathbf{R})$ by choice of σ^2 yields

$$\sigma_o^2 = \frac{\sum_{i=1}^{M} x_i^2 + \sum_{j=1}^{N} y_j^2}{M + N} . \tag{6}$$

This quantity is the sample variance of all the available samples under hypothesis H_o. Maximization of $p_1(\mathbf{R})$ by simultaneous choice of m and σ^2 yields

$$m_1 = \max\left\{\frac{1}{M}\sum_{i=1}^{M} x_i, 0\right\} , \tag{7A}$$

$$\sigma_1^2 = \frac{\sum_{i=1}^{M} (x_i - m_1)^2 + \sum_{j=1}^{N} y_j^2}{M + N} . \tag{7B}$$

Here m_1 is the sample mean of the M signal-plus-noise samples (if non-negative) and σ_1^2 is the sample variance of all available samples under hypothesis H_1 (after subtraction of m_1 from the signal-plus-noise samples). Substitution of (6) and (7) into the LR $p_1(\mathbf{R})/p_o(\mathbf{R})$ yields

$$\mathrm{GLR}(\mathbf{R}) = \begin{Bmatrix} (\sigma_o/\sigma_1)^{M+N} & \text{if } \sum_{i=1}^{M} x_i \geq 0 \\ 1 & \text{otherwise} \end{Bmatrix} . \tag{8}$$

Comparison of the GLR with a threshold T thus becomes merely a comparison of the sample variances under the two hypotheses. The GLR test is therefore to

$$\text{choose } H_1 \text{ if } \left(\frac{\sigma_o}{\sigma_1}\right)^{M+N} > T(\geq 1) \text{ and if } \sum_{i=1}^{M} x_i \geq 0 ; \tag{9}$$

otherwise, choose H_o. By manipulating (6), (7), and (9), the GLR test may be put in the form

$$\sum_{i=1}^{M} x_i \gtrless r\sqrt{M}\left(\sum_{i=1}^{M} x_i^2 + \sum_{j=1}^{N} y_j^2\right)^{1/2} \quad ; \quad r \geq 0 \ . \tag{10}$$

This is the test quoted in Table 1 earlier. The reason for the choice of scale factor as in (10) is that r need vary only in the range (0, 1). This follows by application of Schwartz's inequality to the left side of (10):

$$\left|\sum_{i=1}^{M} x_i\right| \leq \sqrt{M}\left(\sum_{i=1}^{M} x_i^2\right)^{1/2} \quad \text{for any} \quad \{x_i\}_1^M \ . \tag{11}$$

The form of the GLR test in (10) is recognized as a comparison of the sample mean of the signal-plus-noise samples with a scaled version of the sample root-mean-square (rms) value of all the samples. It is a slight generalization of the one-sided t-test [Ref. 1, p. 320]. The test (10) could be adopted even if signal component means on the M samples were unequal, because better processing than (10) provides would require more information than is generally available. However, the present analysis does not cover this more general case of unequal means (see Appendix A).

For case (c), where only σ^2 is unknown, the ML estimates are given again by (6) and (7B) except than m_1 is replaced by the known value m. The GLR test takes the form

$$\left(\frac{\sigma_0}{\sigma_1}\right)^{M+N} \gtrless T \ , \tag{12}$$

which can be manipulated into the form

$$\frac{1}{M}\sum_{i=1}^{M} x_i \gtrless \frac{m}{2} + \frac{\alpha}{m}\left(\sum_{i=1}^{M} x_i^2 + \sum_{j=1}^{N} y_j^2\right). \tag{13}$$

This test requies knowledge of the value of m, as noted earlier, and therefore is discarded.

The derivation of the performance of the GLR test (10) is lengthy and is carried out in a series of appendixes. In Appendix B the matrix manipulations that are needed are reviewed. In Appendix C the method of converting a Hermitian form to a sum of squares of uncorrelated random variables (RVs) is given, along with the statistics of the Hermitian form for Gaussian RVs. In Appendix D the actual detection probability of the GLR test is evaluated and found to be

$$P_D = \int_0^\infty dw \frac{w^K \exp(-w^2/2)}{2^{(K-1)/2}\,\Gamma\left(\frac{K+1}{2}\right)} \Phi\left(d_T - \frac{r}{\sqrt{1-r^2}} w\right) \equiv f(d_T, r, K), \tag{14}$$

where*

$$d_T = \sqrt{M}\,\frac{m}{\sigma}, \quad K = M + N - 2. \tag{15}$$

The quantity d_T can be interpreted as the (voltage) SNR of the RV on the left side of the GLR test. That is, defining

$$z = \sum_{i=1}^{M} x_i, \tag{16}$$

it is easily shown from (2) and (3) that

$$\frac{\text{change in mean of z due to signal presence}}{\text{standard deviation of z}} = \sqrt{M}\,\frac{m}{\sigma} = d_T. \tag{17}$$

Thus, 20 log d_T can be interpreted as a system <u>output</u> SNR in dB; it can be converted to an equivalent input SNR when the receiver processing and filtering operations are known.

---

*If $M = 1$, then $N \geq 1$ is necessary in the GLR test (10) in order to have a valid comparison. However, if $M \geq 2$, then $N \geq 0$ is acceptable. These cases are summarized by requiring $M \geq 1$ and $K \geq 0$ for a meaningful problem.

The scale factor r in the GLR test (10) is under our control, taking values in the range (0, 1), and is adjusted to realize a prescribed false-alarm probability. The false-alarm probability is obtained by setting the SNR $d_T = 0$ in (14):

$$P_F = \int_0^\infty dw \frac{w^K \exp(-w^2/2)}{2^{(K-1)/2}\Gamma\left(\frac{K+1}{2}\right)} \Phi\left(-\frac{r}{\sqrt{1-r^2}}w\right) = f(0, r, K) . \tag{18}$$

In Appendix E, closed-form expressions for P_F and simple recursive relations for P_D are derived. These are used in the following section to obtain numerical results.

It will be noticed in the GLR test (10) that the signal-plus-noise samples $\{x_i\}$ are inbred; that is, they are used on both sides of the comparison. It is of interest to compare this with the test that uses the samples, $\{x_i\}$, only according to

$$\sum_{i=1}^{M} x_i \gtrless \alpha \left(\sum_{j=1}^{N} y_j^2\right)^{1/2} ; \quad \alpha \geq 0 . \tag{19}$$

The performance for (19) is easily determined when we compare it with (D-27): test (19) is already in the desired form of a comparison of independent RVs; however, there are $M + N - 1$ RVs on the right side of the comparison in (D-27), whereas there are only N RVs in (19). Since the addition of more zero-mean RVs can only help give better estimates of the rms noise level, the performance of the GLR test (10) must be better than the performance of test (19).

As N tends to infinity in (10), the estimate of the rms noise level on the right side becomes perfect. Test (10) then becomes a comparison with a constant, and the detection and false-alarm probabilities are then

$$P_D = \Phi(d_T - \Lambda), \quad P_F = \Phi(-\Lambda) . \tag{20}$$

These are the values approached by (14) and (18), respectively, as K tends to infinity and r tends to zero in such a way that a specified P_F is realized (by choice of Λ).

RESULTS

It will be seen from (18) that the only way that M and N enter P_F is through their sum; that is, $K = M + N - 2$. This is due to the way the GLR test was set up in (10); it will be noticed that M enters the test separately as a scaling factor. For $P_F = 10^{-n}$, $n = 1(1)8$, and $K = 0(1)10(5)25$, required values of r in the test (10) are presented in Table 2.

In Figs. 1-14, P_D is plotted versus d_T, with P_F as a parameter, for $K = 0(1)10(5)25$. The curve for $K = \infty$, as discussed with respect to (20), is presented in Fig. 15 for comparison purposes. It is immediately obvious from the curve for $K = 0$ that performance for $M = 2$, $N = 0$ or $M = 1$, $N = 1$ is very poor. Inadequate estimation of the noise level requires a high threshold-level setting for low P_F and thereby forces down the attainable values of P_D. In fact, even for $d_T = 20$, P_D can not reach a value of 0.5 for $P_F \leq 10^{-2}$. The situation for moderate P_F is quickly improved by the addition of a few extra samples, as the curves for $K = 1$, 2, and 3 show. However, high-quality performance is not realized for very small P_F until many samples are available; for example, even for $d_T = 20$, P_D can not reach a value of 0.5 for $P_F = 10^{-8}$ until $K \geq 8$.

The addition of another independent sample always improves performance; however, the improvement is greater if M, rather than N, is increased by one. This may be seen by noting that increasing N by one increases K by one, whereas increasing M by one increases both K and d_T. Thus, an additional potential signal sample is always preferable to a noise-alone sample.

The improvement to be gained by collecting additional samples is more obvious in Figs. 16 and 17, which are plots of the values of d_T versus K required to realize $P_D = 0.5$ for various values of P_F. Figure 17 shows d_T in dB, that is, $20 \log d_T$; straight lines have been drawn between points for ease of interpretation. The horizontal tic marks at the right edges of the figures indicate the required value of d_T for $K = \infty$ and are therefore the asymptotes of the curves.

For $P_F = 10^{-1}$, the curve has almost reached its limiting value at $K = 1$ or 2. For $P_F = 10^{-2}$, the required value of d_T for $K = 0$ is 21.5, but the required value of d_T for $K = 1$ is 5.83, a drop of 11.3 dB, which is a marked improvement for the addition of one more sample. For $P_F = 10^{-8}$, however, the curve is still dropping at $K = 25$.

Table 2

REQUIRED VALUES OF SCALE FACTOR r IN GLR TEST (EQ. 10)

K	P_F							
	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
0	.9511	.9995	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.8000	.9800	.9980	.9998	1.0000	1.0000	1.0000	1.0000
2	.6870	.9343	.9859	.9970	.9993	.9999	1.0000	1.0000
3	.6084	.8822	.9633	.9884	.9963	.9988	.9996	.9999
4	.5509	.8329	.9350	.9743	.9898	.9960	.9984	.9994
5	.5067	.7887	.9049	.9564	.9799	.9907	.9957	.9980
6	.4716	.7498	.8751	.9365	.9674	.9832	.9913	.9955
7	.4428	.7155	.8467	.9156	.9531	.9738	.9853	.9918
8	.4187	.6851	.8199	.8947	.9378	.9630	.9779	.9868
9	.3981	.6581	.7950	.8742	.9219	.9512	.9694	.9808
10	.3802	.6339	.7717	.8544	.9060	.9389	.9601	.9739
15	.3170	.5425	.6777	.7676	.8303	.8751	.9076	.9313
20	.2774	.4815	.6099	.7001	.7665	.8167	.8553	.8853
25	.2497	.4372	.5587	.6468	.7138	.7662	.8080	.8417

If numerous signal samples are already available, such as $M > 10$, Figs. 16 and 17 indicate that it is not worthwhile to collect any noise-alone samples unless a very small P_F is desired. For example, the curve for $P_F = 10^{-3}$ in Fig. 17 indicates that the required value of d_T would drop less than 2 dB by increasing K from 8 to 25, which is within one dB of the $K = \infty$ case. However, at $P_F = 10^{-8}$, a 7-dB reduction is possible by increasing K from 8 to 25.

The total received signal energy is proportional to Mm^2. If this total energy is kept fixed, dividing it into more components by increasing M will increase K while d_T remains constant. In Figs. 1-14, it is seen that performance improves monotonically; that is, the more the total signal energy is fractionalized, the better the performance approaches the $K = \infty$ limit shown in Fig. 15. This is due to the fact that coherent addition of the signal components is presumed (even at low SNR per component), while the noise-power estimate becomes more stable. In practice, the required coherent processing may be unattainable for low SNRs.

A word of caution is in order regarding the use of Figs. 1-17. The quantity d_T depends on M, as indicated in (17). Thus, if M changes, both K and d_T change; however, if N changes, K alone changes. The plots presented here can be used, for example, to compute individual SNR m/σ curves for any values of M and N, but, in order to keep the number of curves down to a reasonable level, this procedure was not adopted. Notice that the case of no noise-alone samples, $N = 0$, is subsumed under the above results by setting $K = M - 2$.

Curves of required SNR for other detection probabilities, such as 0.9 or 0.99, can be easily determined from Figs. 1-15. Operating characteristics for other ranges of K or SNR can be determined from the general formulas in Appendix E.

DISCUSSION

Operating characteristics for detection of coherent signals in Gaussian noise of unknown level have been presented. The processor considered is the ML detector in the case of unknown signal level. It is the optimum processor in the following sense: out of the (infinite) class of processors that yield a specified P_F without knowledge of the noise level, it yields the maximum P_D for known signal level and is UMP for unknown (positive) signal level. The scaling-invariance optimum with respect to unknown noise level generally yields

processors that are not UMP with respect to the parameters in the PDFs under the two hypotheses H_0 and H_1. Thus one must generally resort to a different approach, such as the principle of ML, that eliminates unknown parameters.

If the polarity of the mean m in the PDF (3) is not known to be nonnegative, the least favorable situation occurs when positive and negative values occur equally often. Then the average LR test takes the form of (10), except that the left side is replaced by its magnitude. (This test is the same as that yielded by the optimum processor forced to operate with the samples $\{x_i/y_N\}$ and $\{y_i/y_N\}$ (see Appendix A, (A-17) - (A-20)).) Derivation of the performance of this processor parallels that given in this report up to (D-27), at which point v_1 is replaced by its magnitude. The detection probability is then given by

$$P_D = \int_0^\infty dw \frac{w^K \exp(-w^2/2)}{2^{(K-1)/2}\,\Gamma\left(\frac{K+1}{2}\right)}\left[\Phi\left(d_T - \frac{r}{\sqrt{1-r^2}}w\right) + \Phi\left(-d_T - \frac{r}{\sqrt{1-r^2}}w\right)\right]$$

$$= f(d_T, r, K) + f(-d_T, r, K)\,. \tag{21}$$

The false-alarm probability is

$$P_F = 2f(0, r, K)\,, \tag{22}$$

which is double expression (18). The performance of this processor is poorer than the one analyzed in this report because of less available knowledge. The exact quantitative behavior may be obtained from (21), (22), and Appendix E, but has not been pursued.

Coherent processing has been presumed here. When the received signal is narrowband, phase tracking is often difficult to accomplish; then phase-incoherent processing is often adopted with an attendant loss of performance. Under certain conditions, such as multipath fading, the received signal possesses no deterministic behavior and is a stochastic process. The two cases of phase-incoherent and stochastic signal statistics will be the subjects of future work.

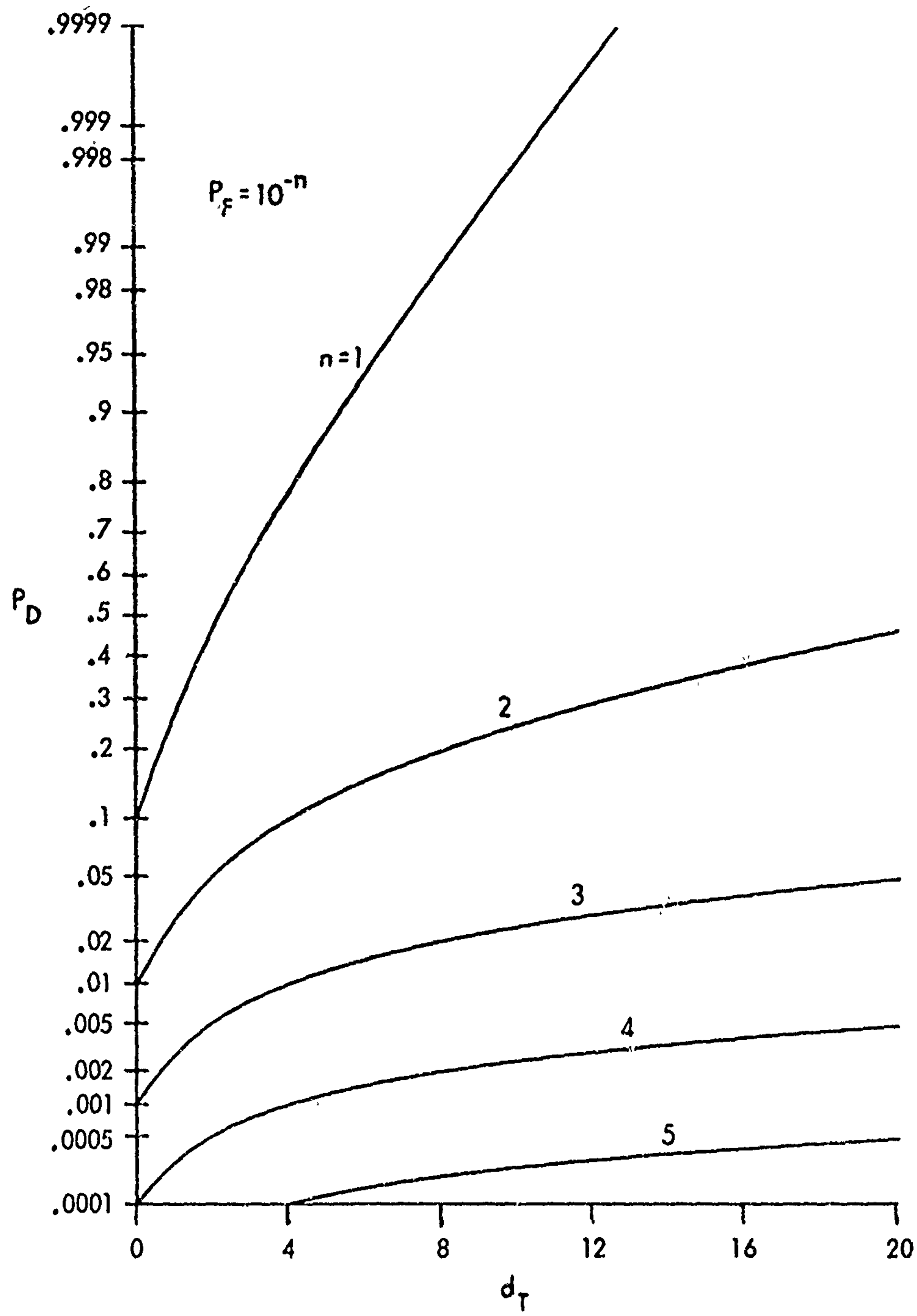

Fig. 1. Detection Probability for K = 0

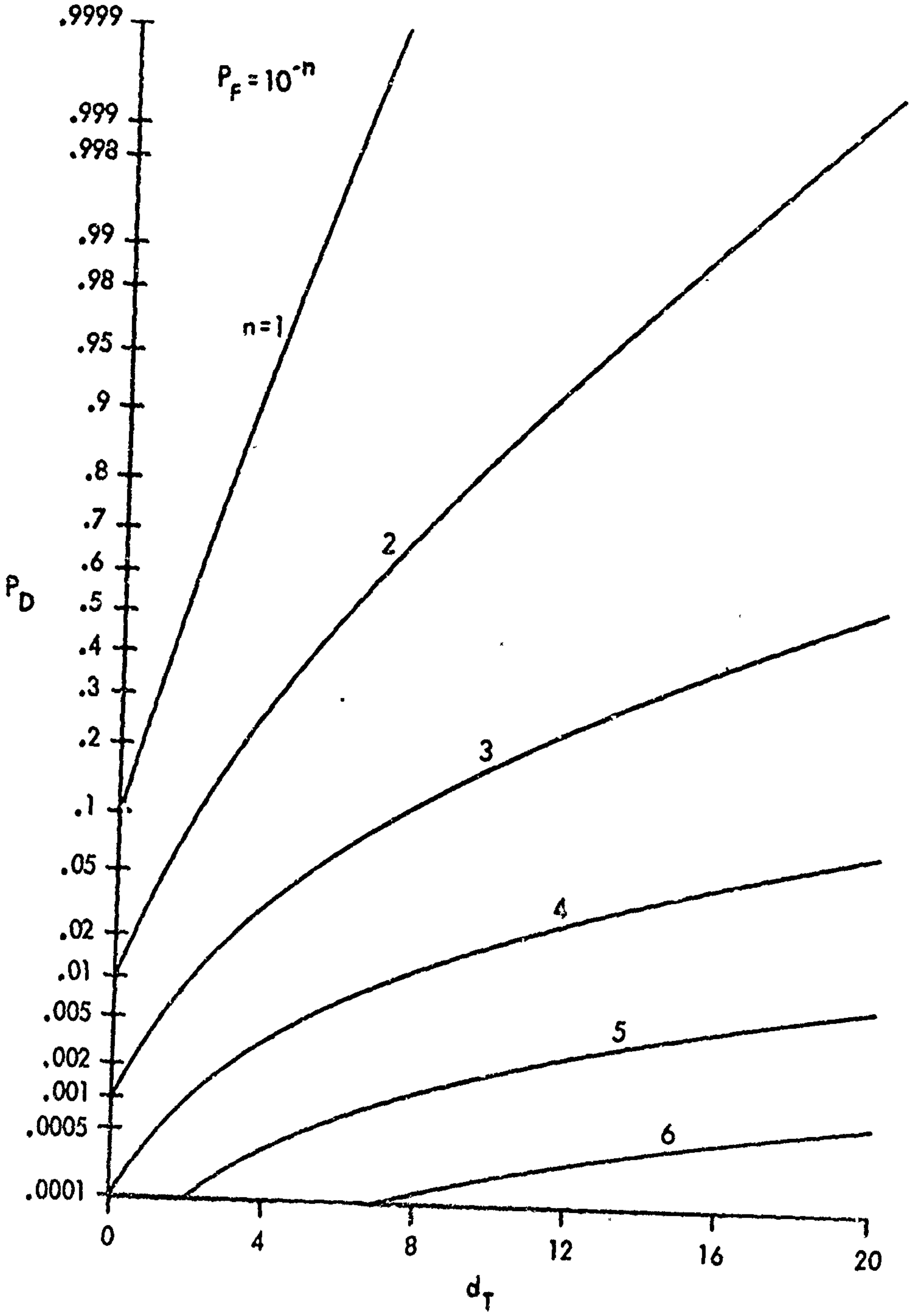

Fig. 2. Detection Probability for K = 1

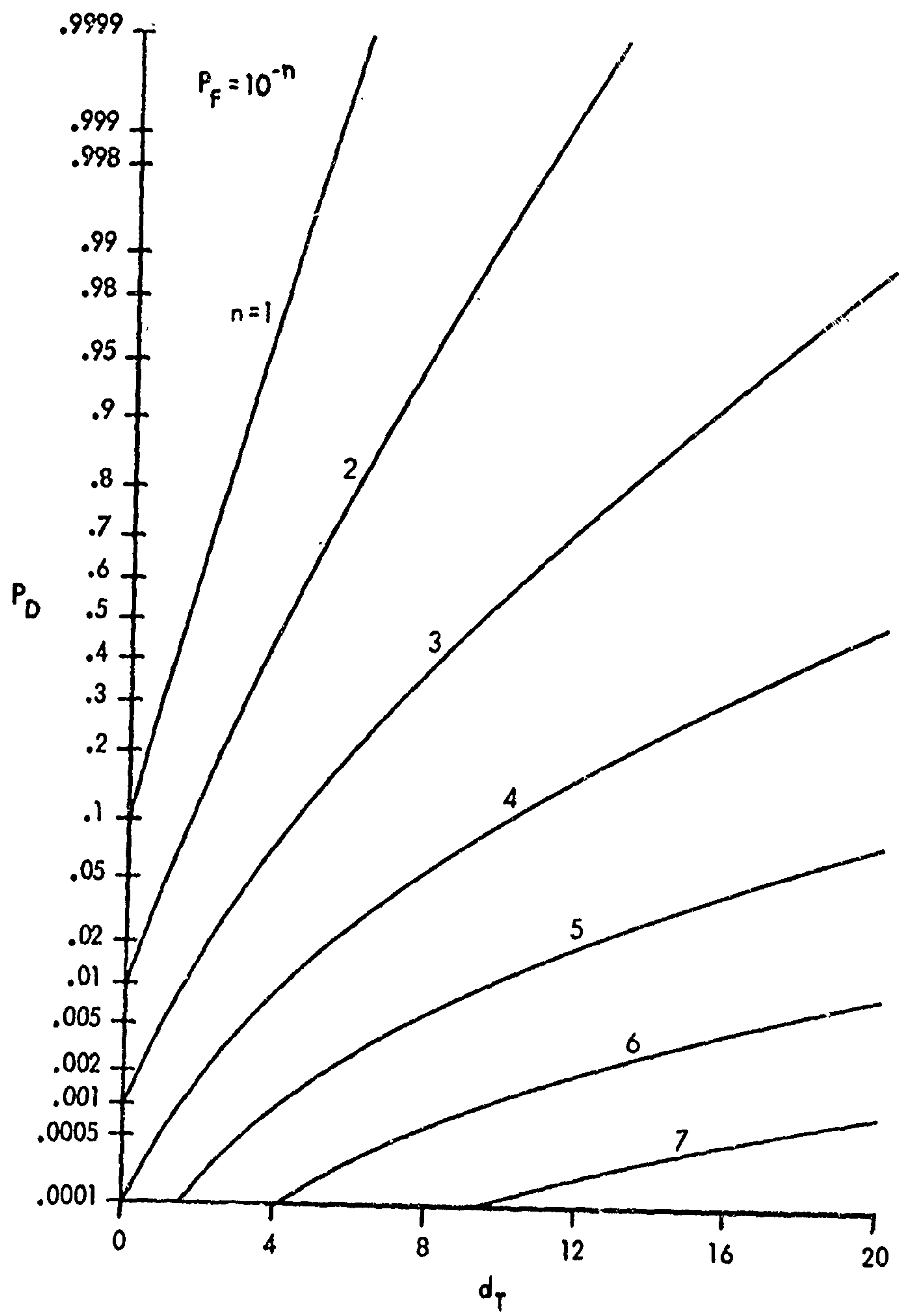

Fig. 3. Detection Probability for K = 2

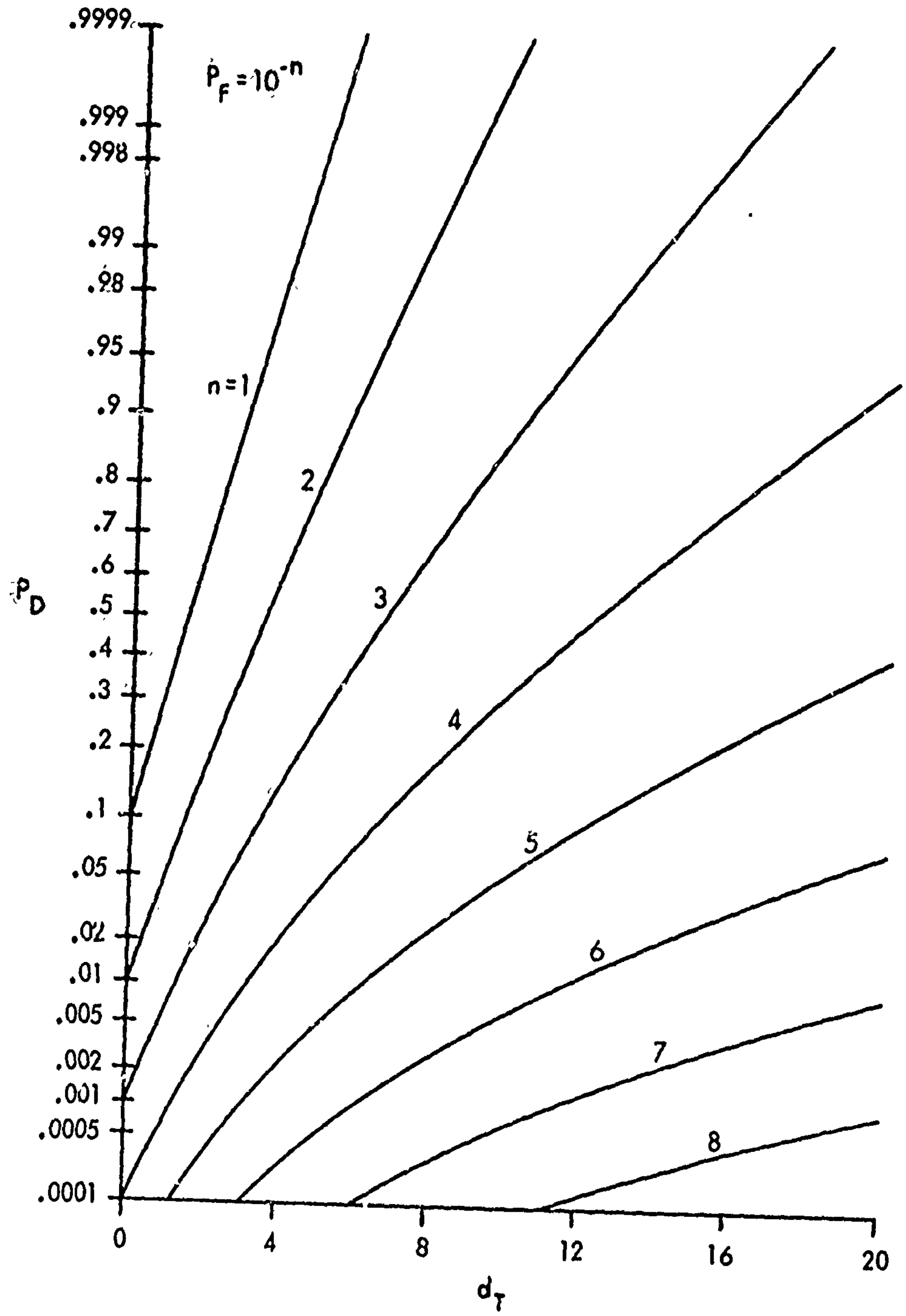

Fig. 4. Detection Probability for K = 3

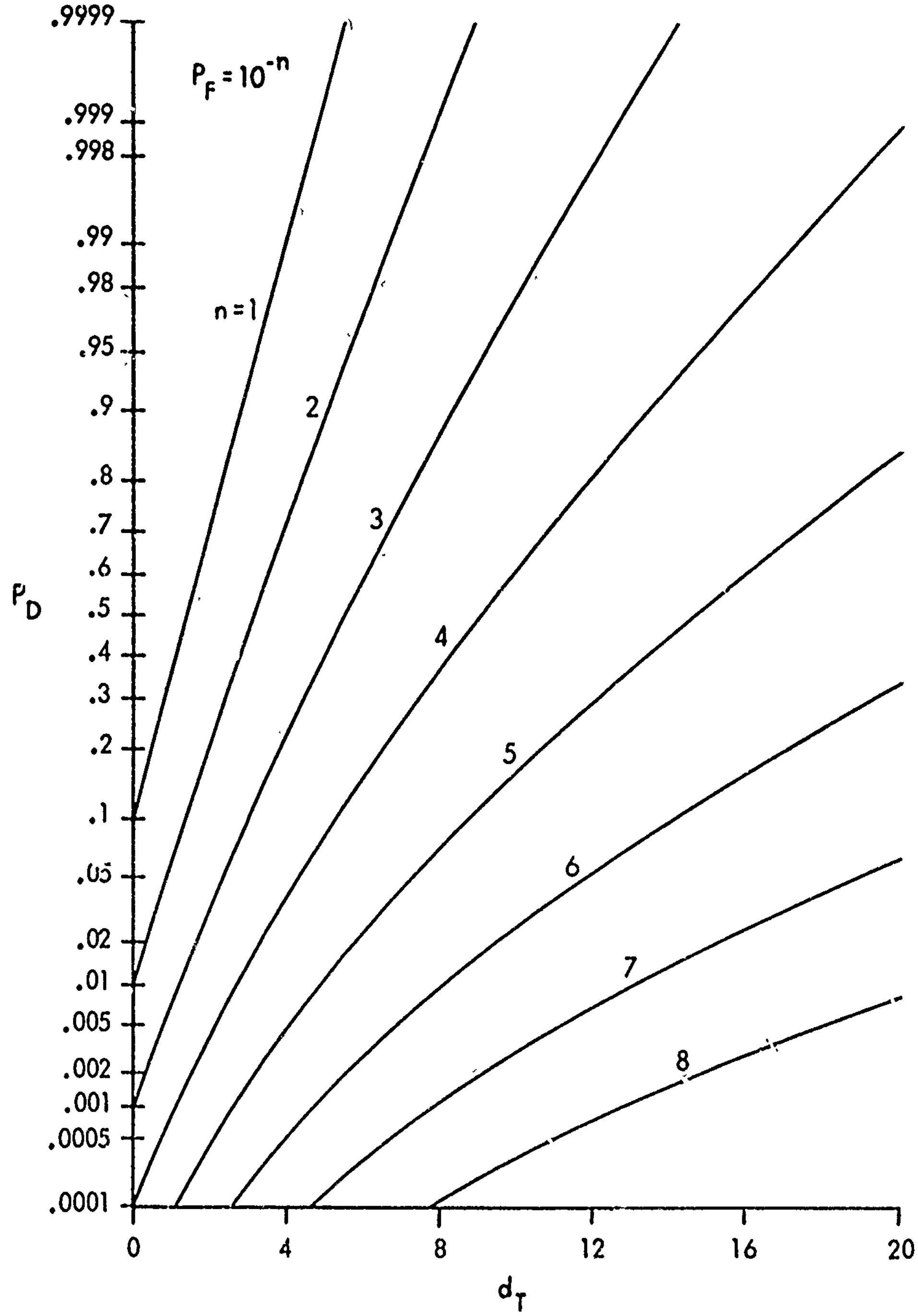

Fig. 5. Detection Probability for K = 4

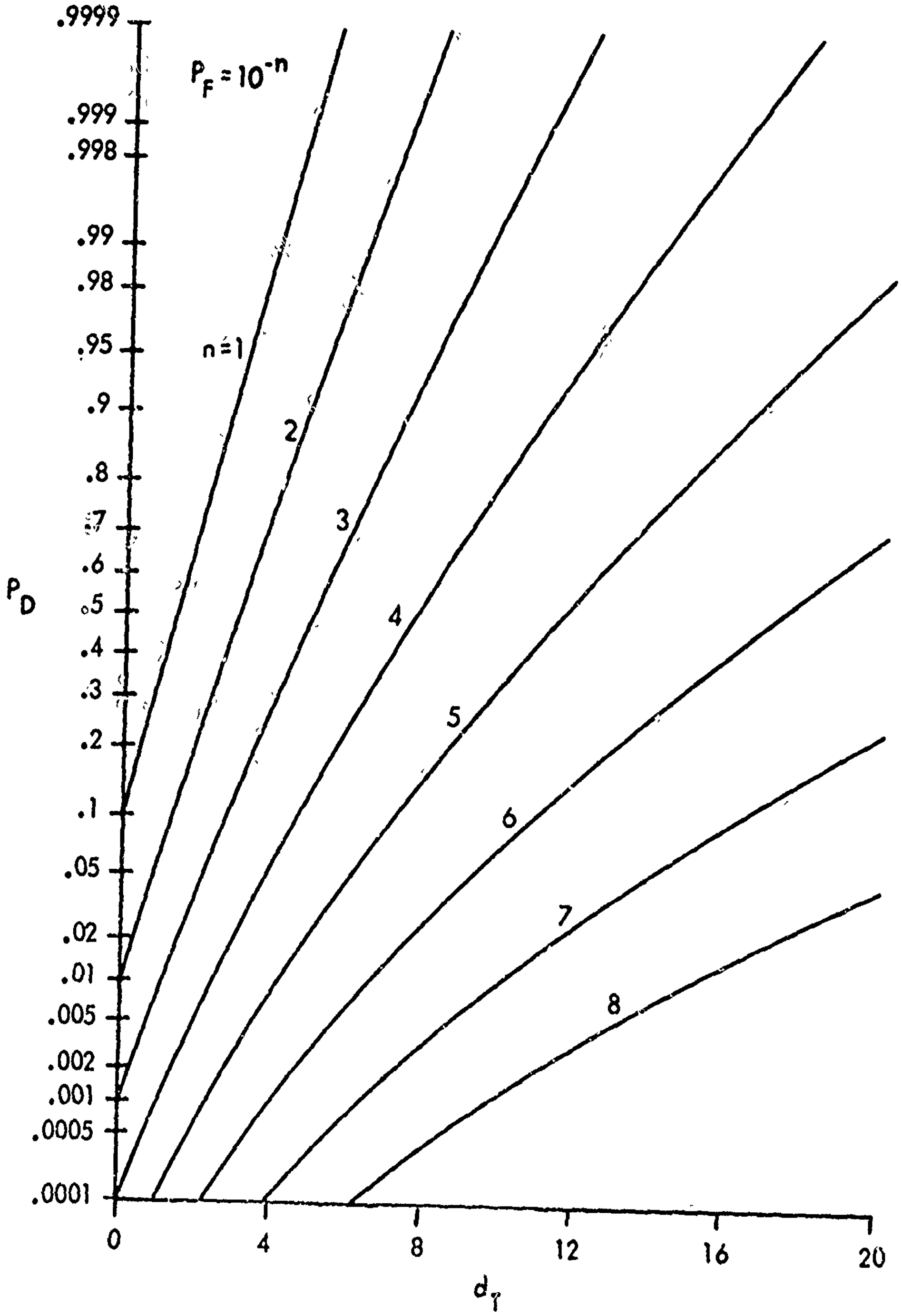

Fig. 6. Detection Probability for K = 5

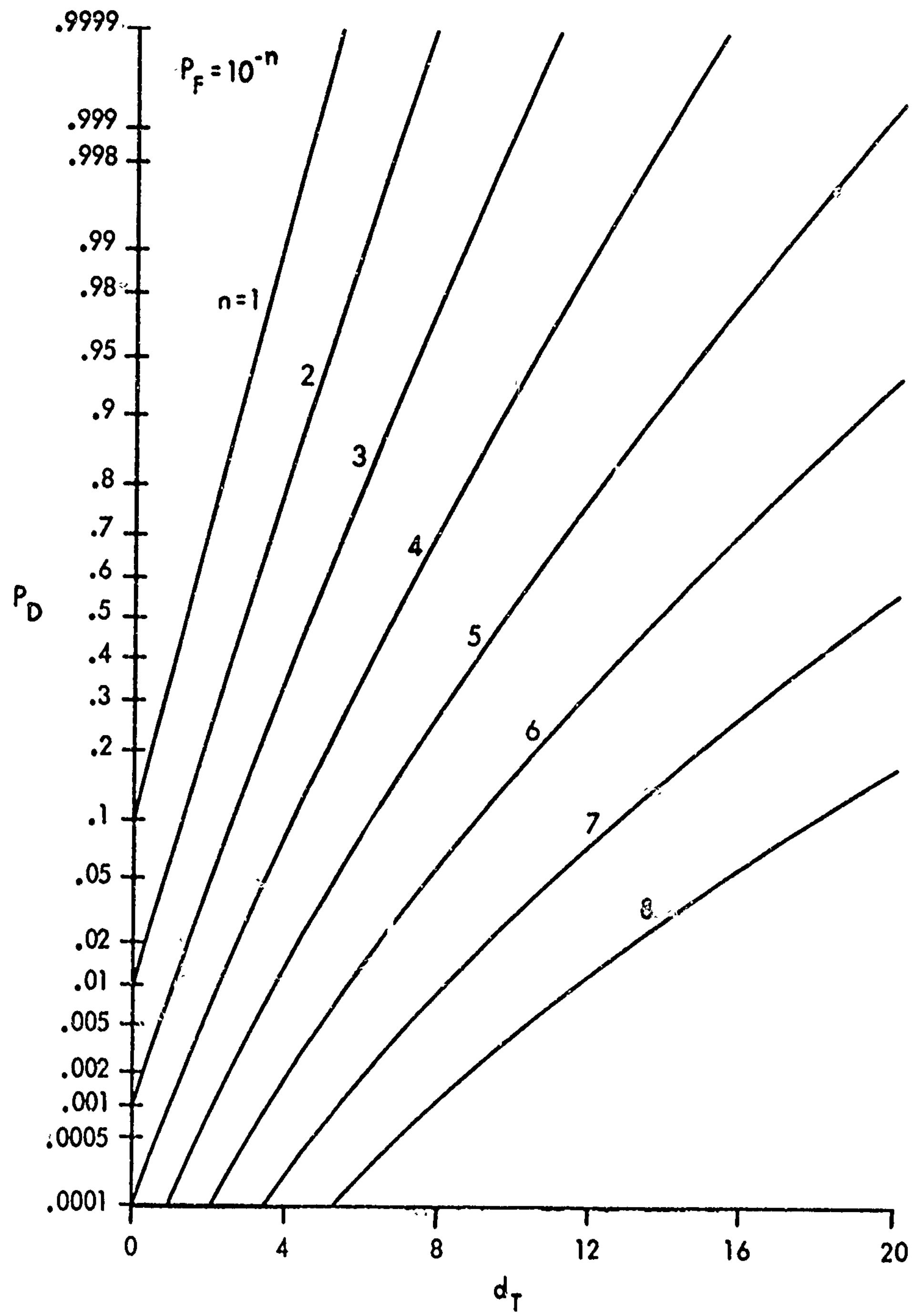

Fig. 7. Detection Probability for K = 6

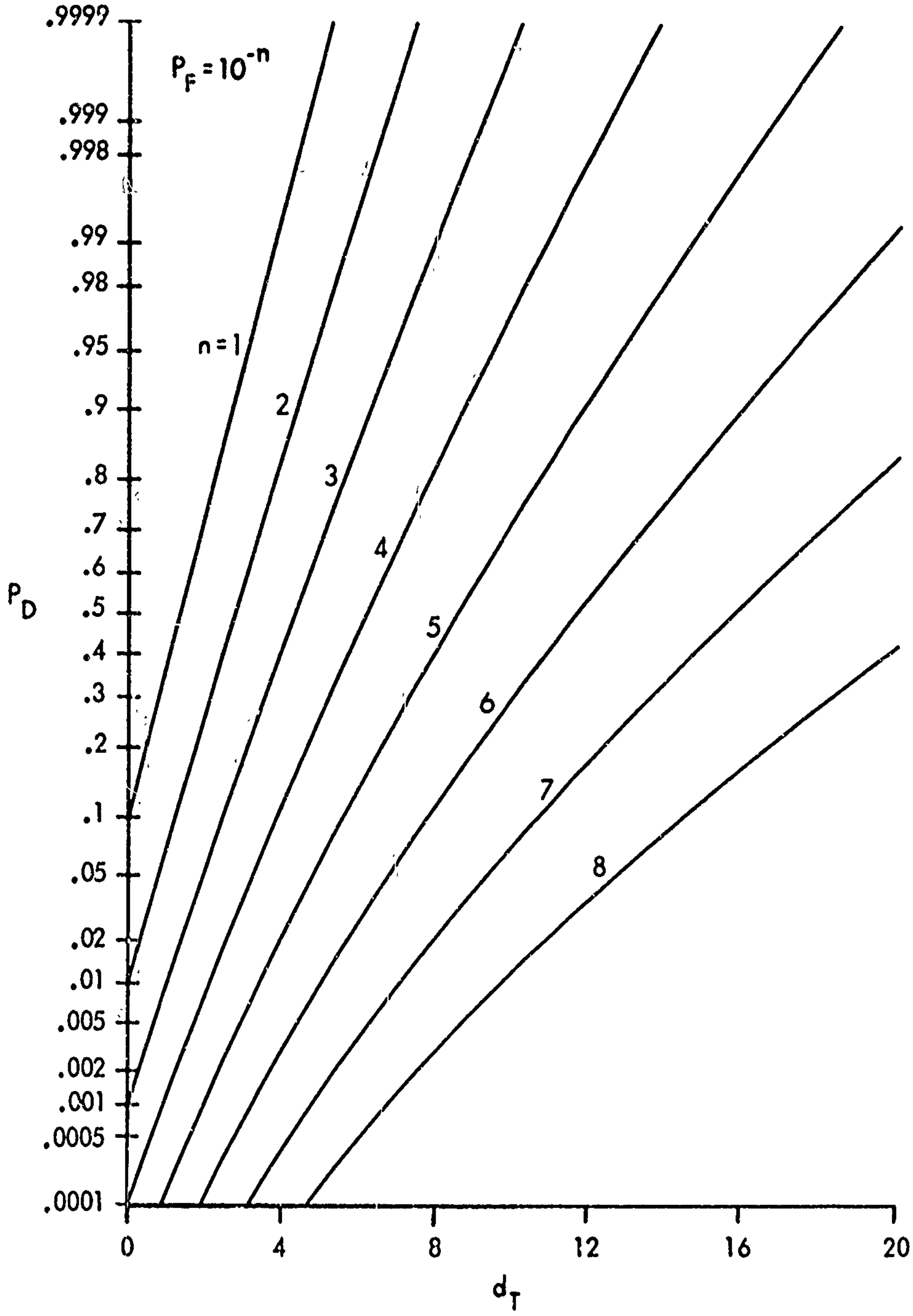

Fig. 8. Detection Probability for K = 7

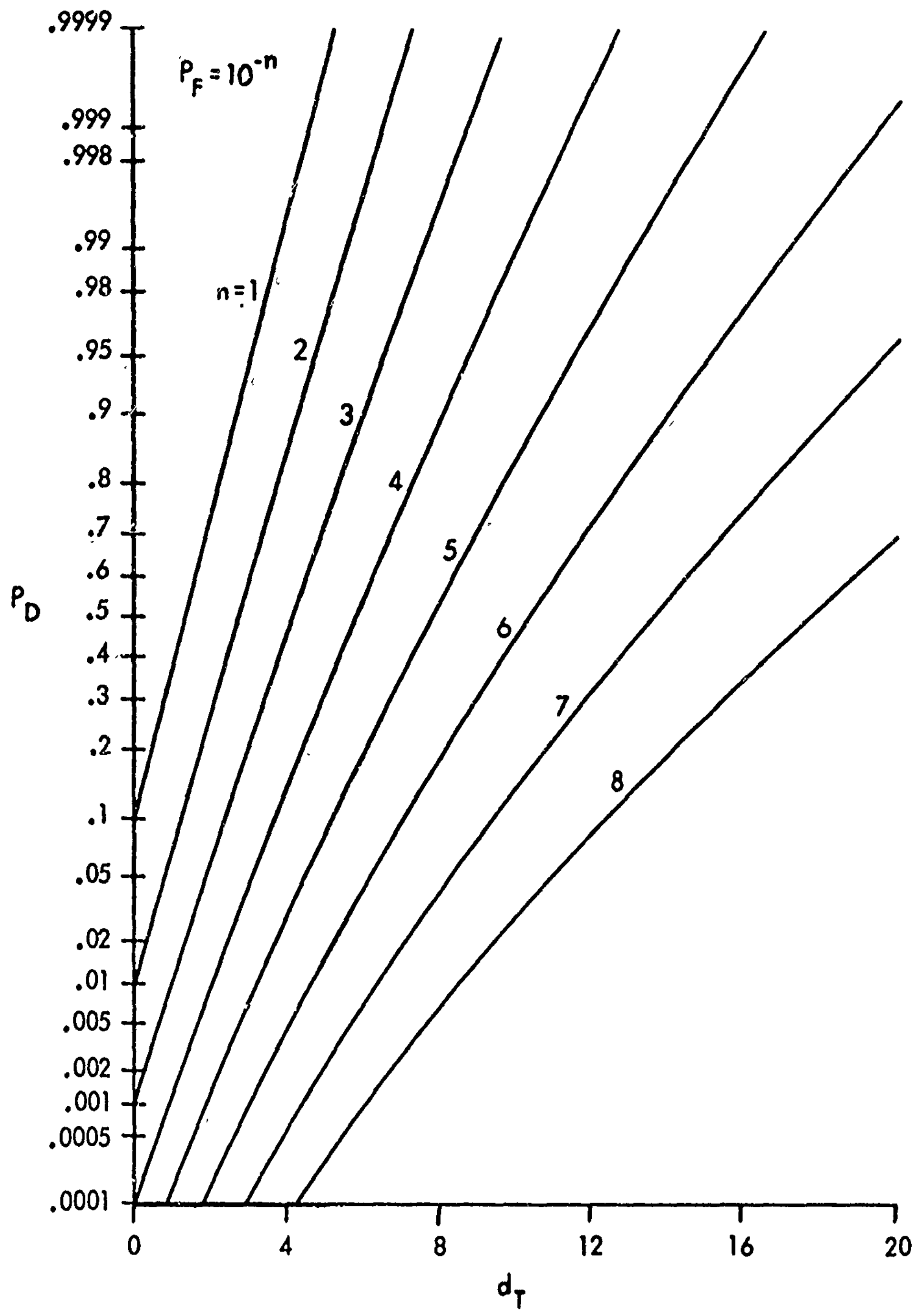

Fig. 9. Detection Probability for K = 8

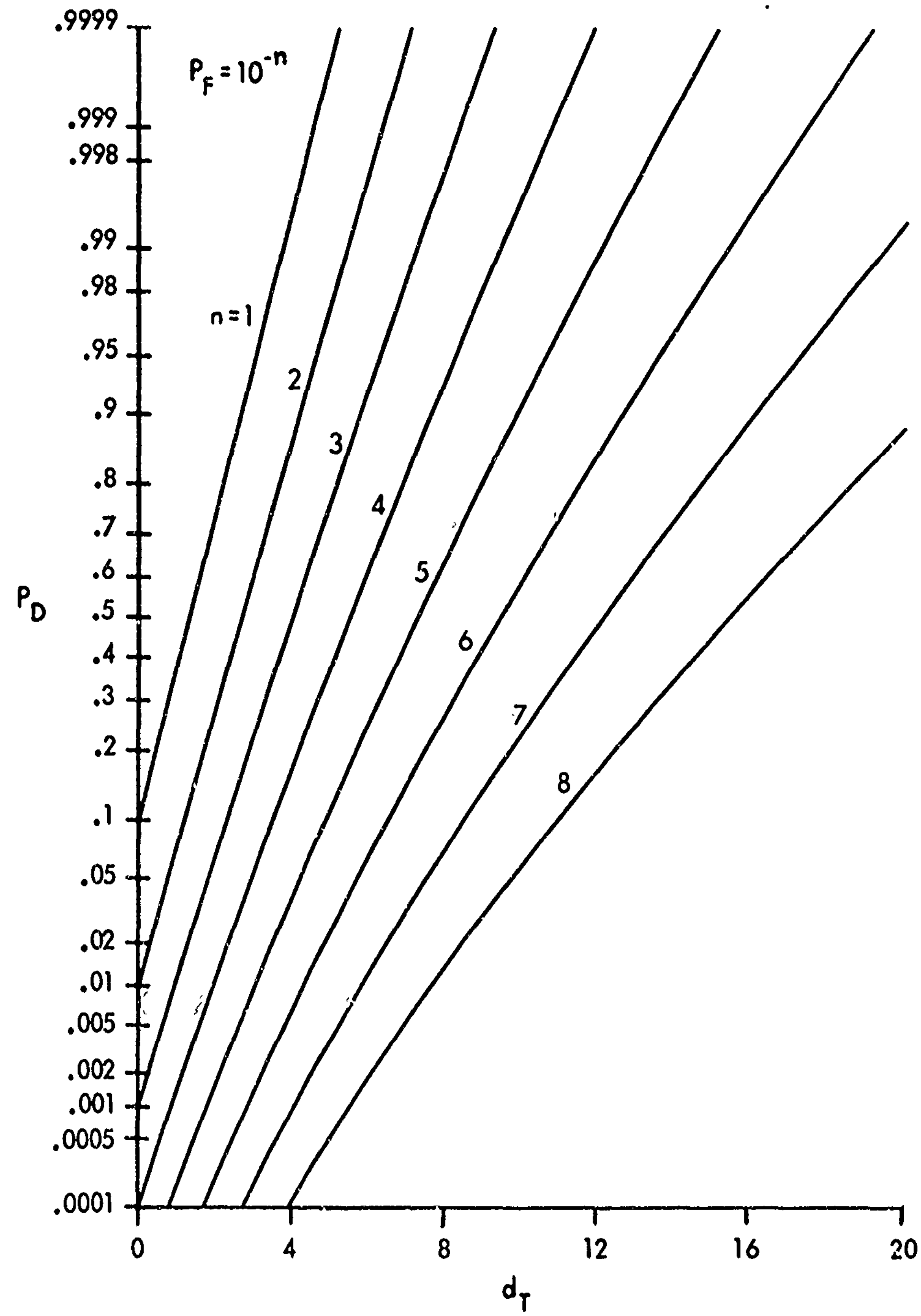

Fig. 10. Detection Probability for K = 9

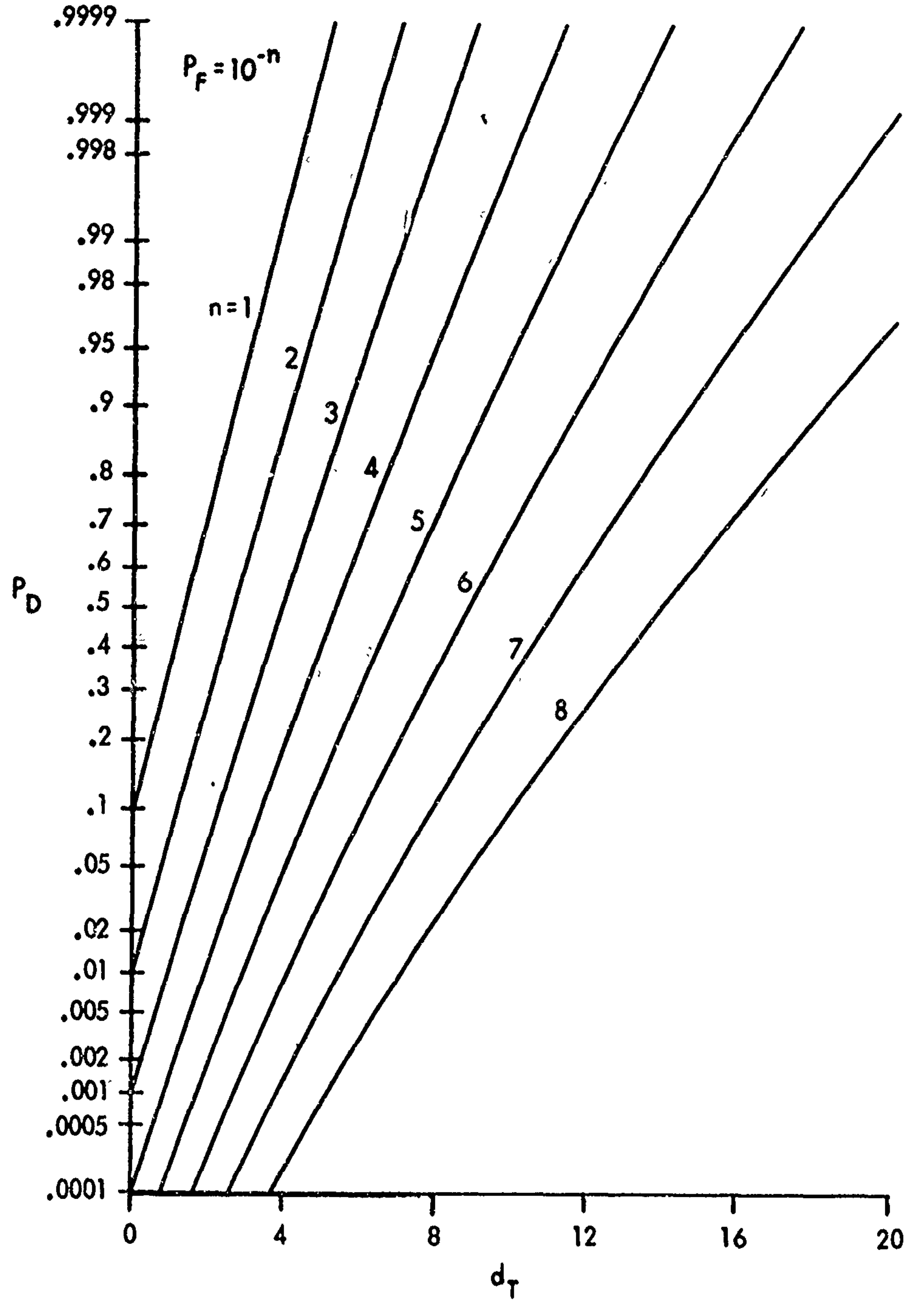

Fig. 11. Detection Probability for K = 10

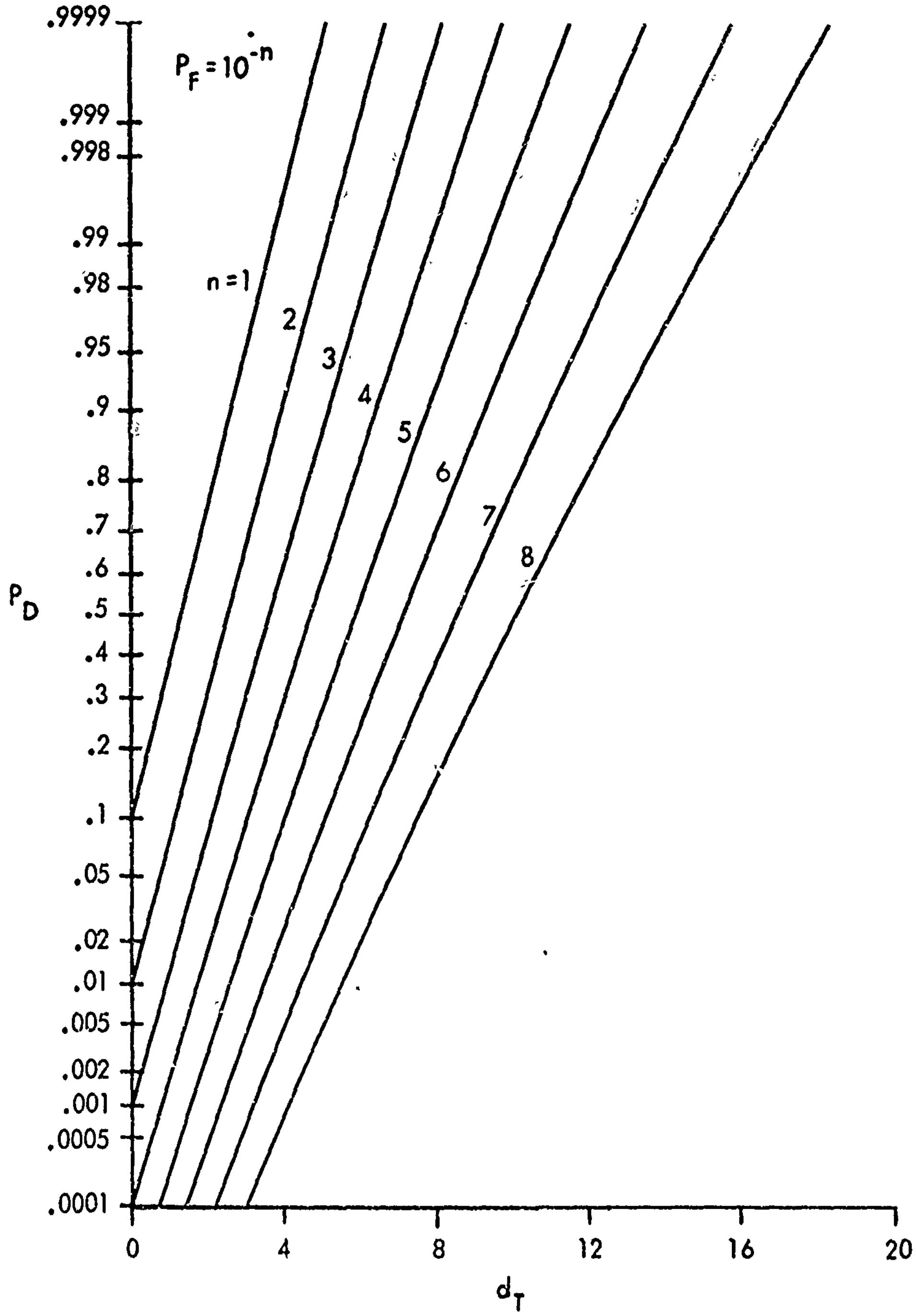

Fig. 12. Detection Probability for K = 15

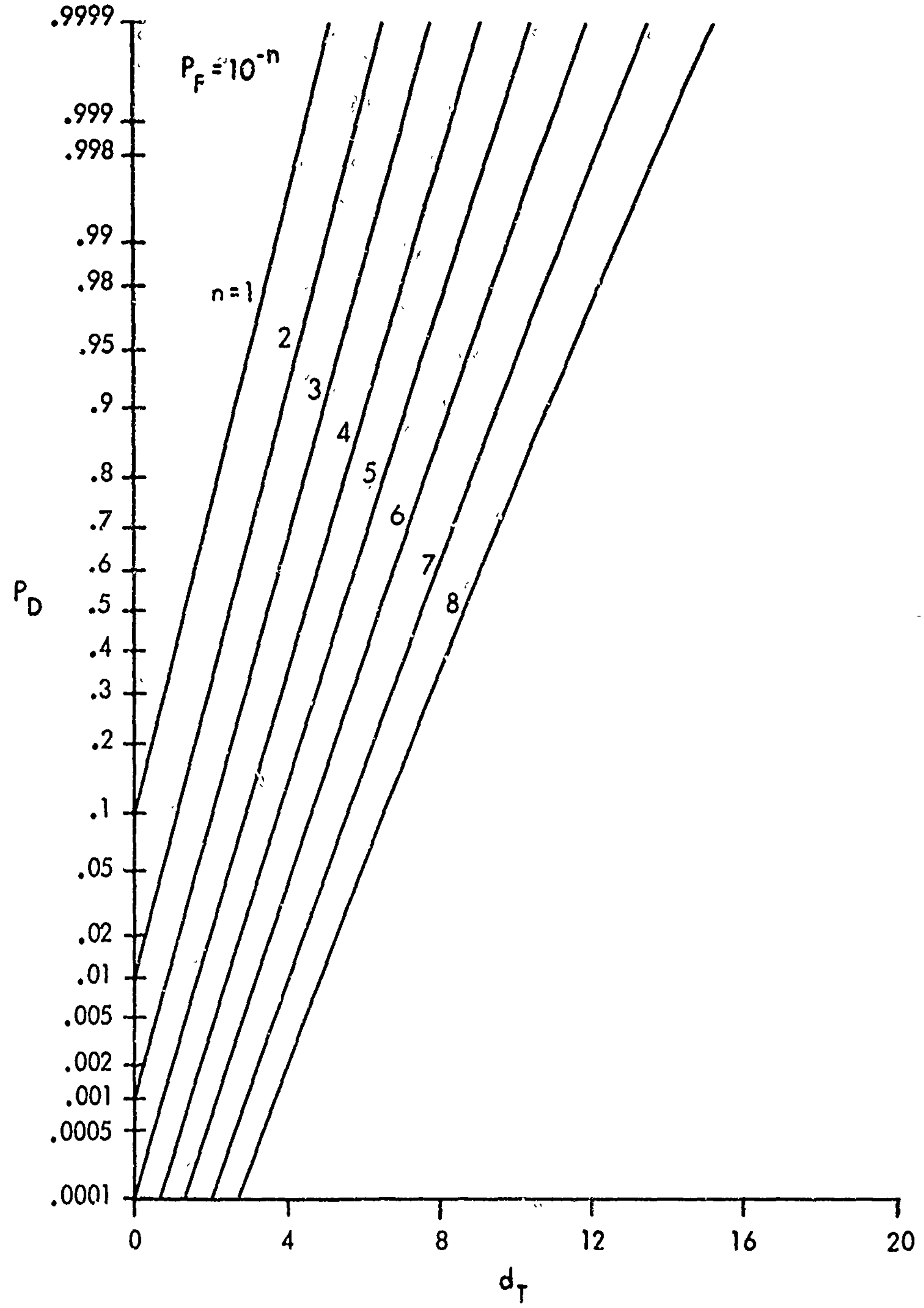

Fig. 13. Detection Probability for K = 20

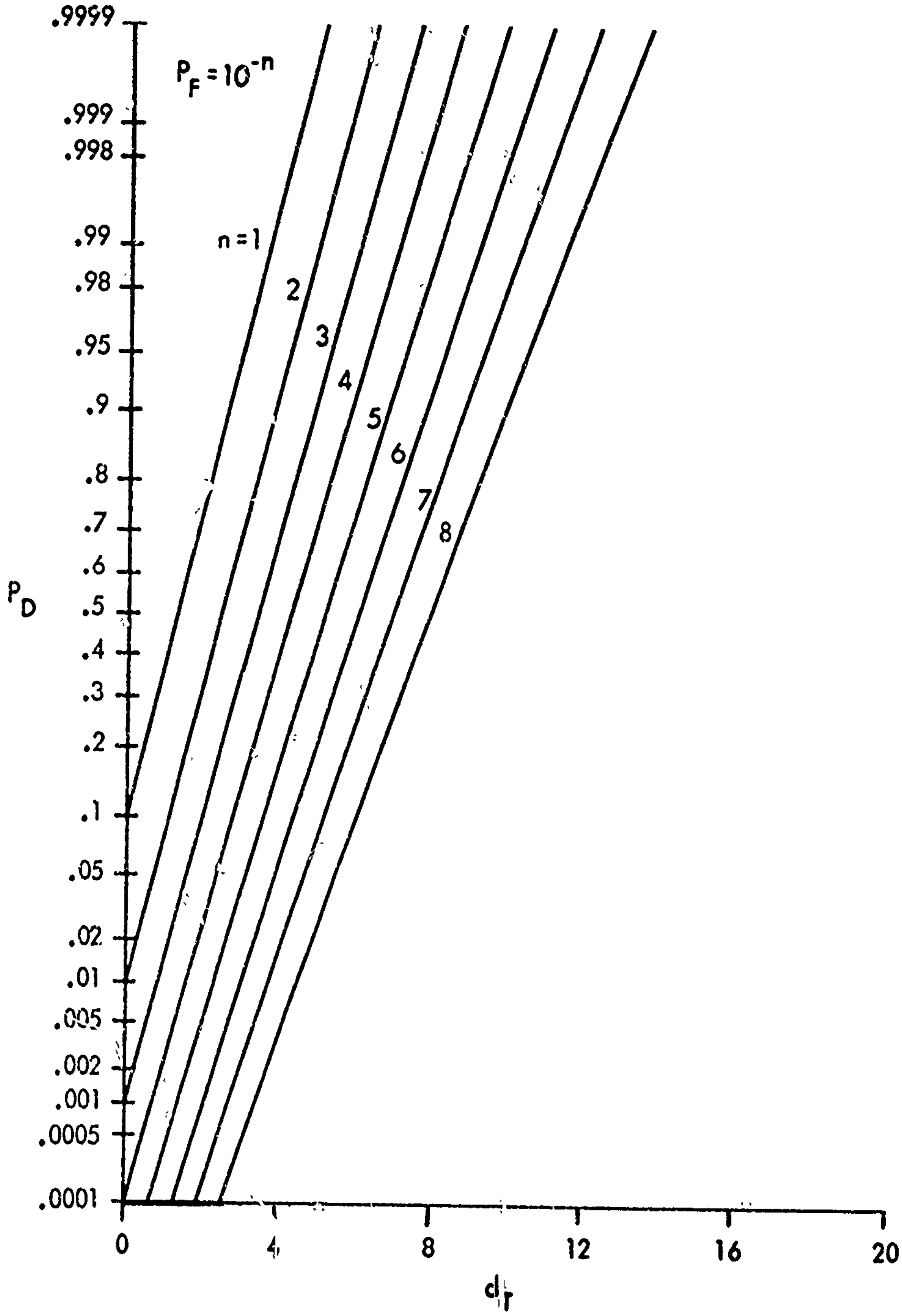

Fig. 14. Detection Probability for K = 25

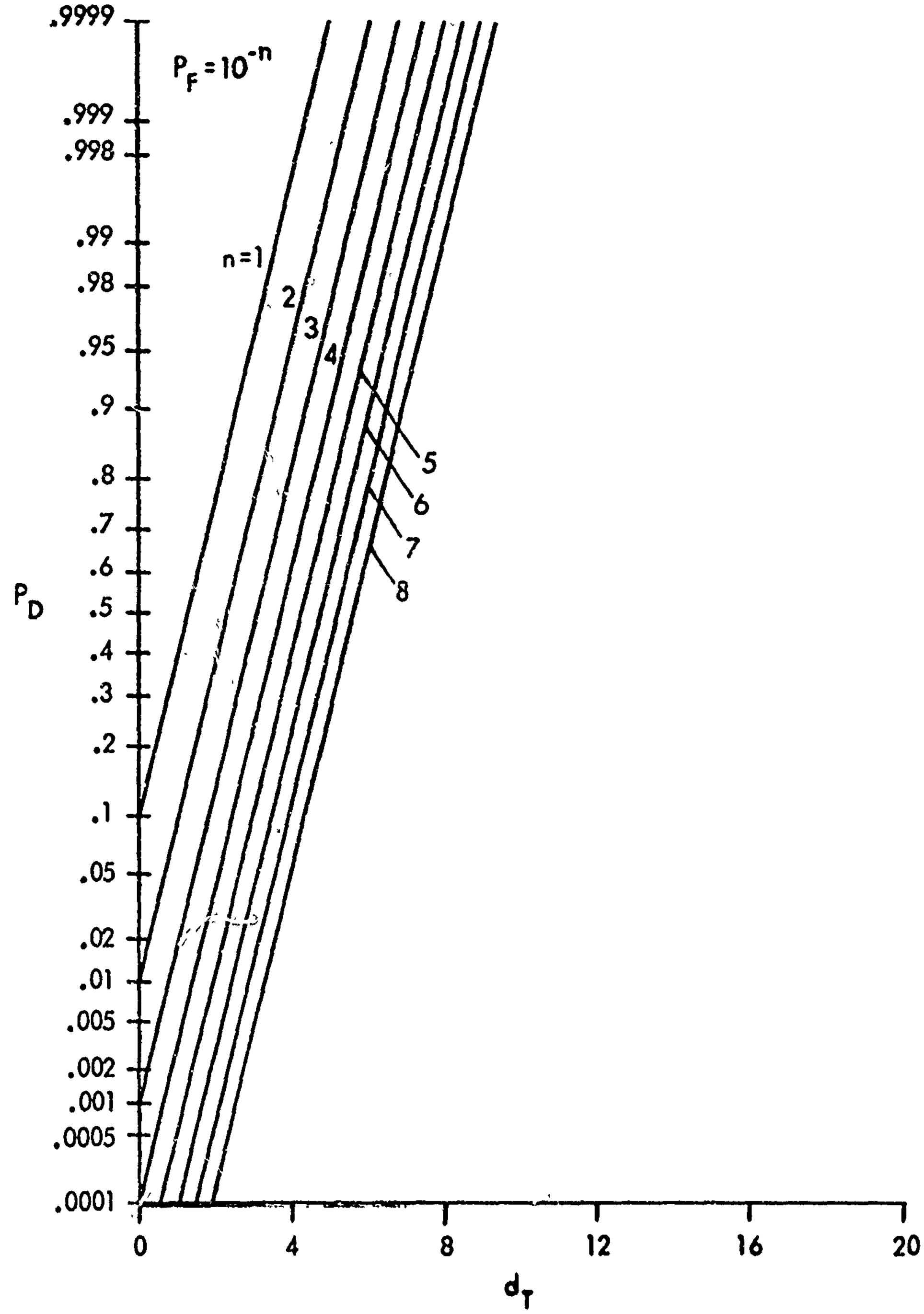

Fig. 15. Detection Probability for K = ∞

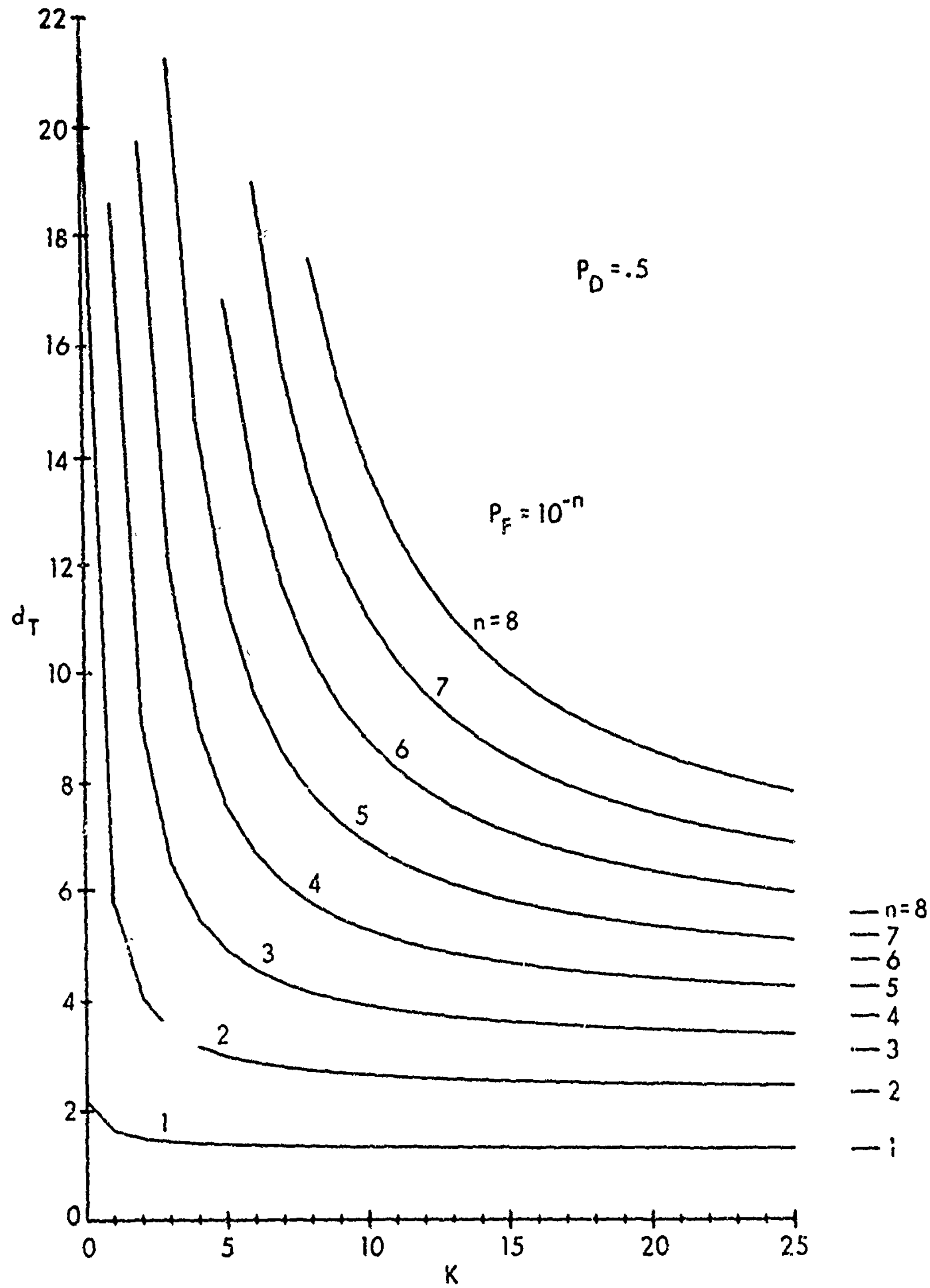

Fig. 16. Required Value of d_T for $P_D = 0.5$

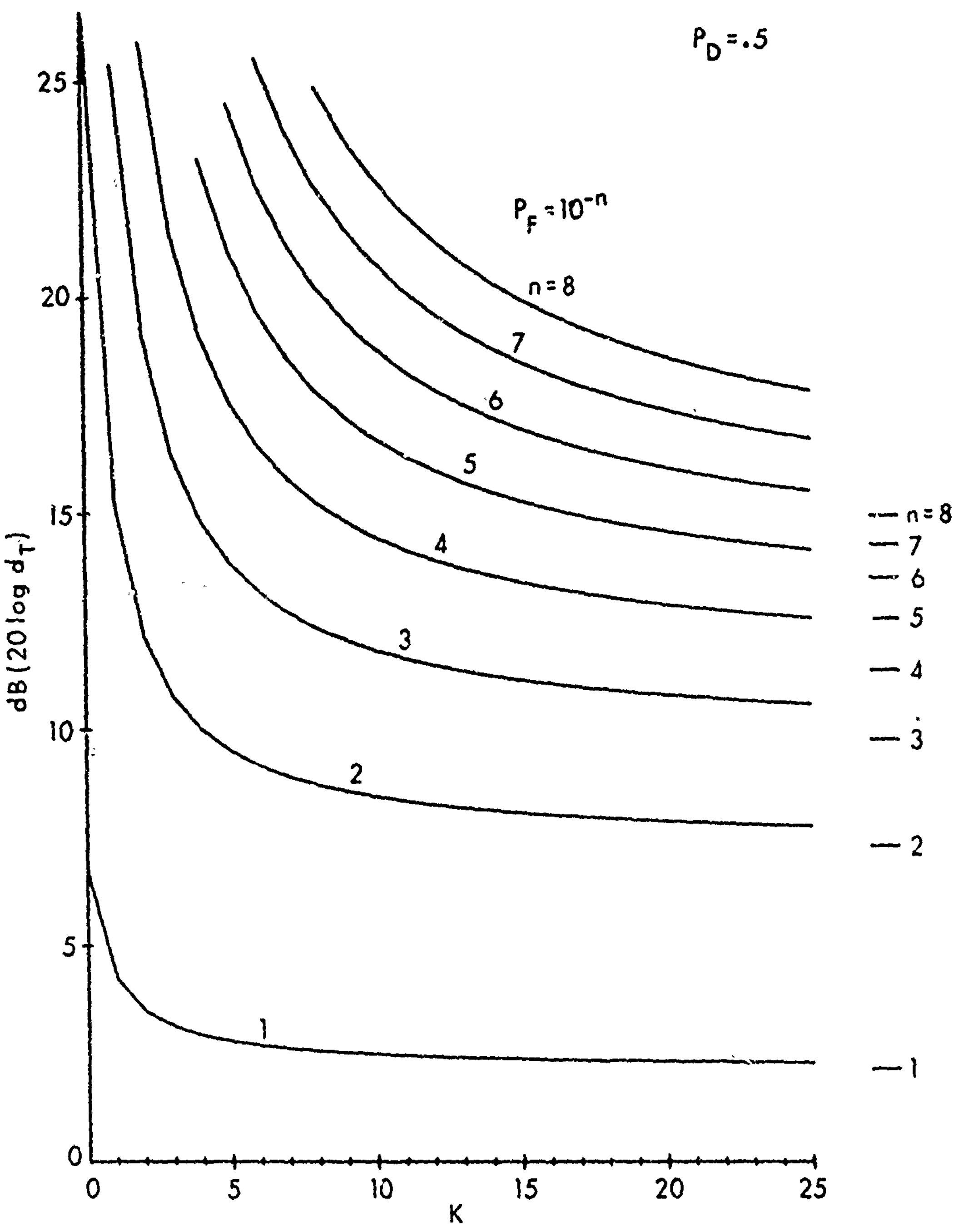

Fig. 17. Required Value of d_T in dB for $P_D = 0.5$

Appendix A

OPTIMUM PROCESSOR FOR ARBITRARY SCALING

Suppose n+1 observations $x_1, \ldots, x_n, x_{n+1}$ are made of a process of unknown noise power level. Let the PDFs of the observations be p and q under hypotheses H_1 and H_0, respectively; the observations need not be statistically independent. We want a processor that yields a specified P_F without knowledge of the noise power level and that realizes maximum P_D. Since the desired processor is to realize a specified P_F without knowledge of the noise level, it must operate equally well on the scaled samples $cx_1, \ldots, cx_n, cx_{n+1}$, where c is any positive constant, since these could just as well have been the sample values under H_0. In particular, we could choose $c = |x_{n+1}|^{-1}$ for the scale factor and still expect the desired processor to perform equally well.

The optimum processor, in the sense of yielding maximum P_D, for the scaled samples

$$z_k = \frac{x_k}{|x_{n+1}|}, \quad 1 \le k \le n+1, \tag{A-1}$$

is one that computes the LR of these samples and compares the LR with a threshold. Noting that $z_{n+1} = \pm 1$, we have

$$LR(z_1, \ldots, z_n | z_{n+1} = \pm 1) \equiv \frac{\overline{p}(z_1, \ldots, z_n | z_{n+1} = \pm 1)}{\overline{q}(z_1, \ldots, z_n | z_{n+1} = \pm 1)}. \tag{A-2}$$

Now, under H_1,

$$\text{Prob}(z_1 < Z_1, \ldots, z_n < Z_n | z_{n+1} = \pm 1) = \text{Prob}(x_1 < |x_{n+1}| Z_1, \ldots, x_n < |x_{n+1}| Z_n | z_{n+1} = \pm 1)$$

$$= \int_{A(z_{n+1})}^{B(z_{n+1})} dx_{m+1} \int_{-\infty}^{|x_{n+1}| Z_1} dx_1 \ldots \int_{-\infty}^{|x_{n+1}| Z_n} dx_n \; p(x_1, \ldots, x_n, x_{n+1}), \tag{A-3}$$

Preceding page blank

where

$$\Big(A(z_{n+1}),\ B(z_{n+1})\Big) = \begin{Bmatrix} (0,\ \infty),\ z_{n+1} = +1 \\ \\ (-\infty, 0),\ z_{n+1} = -1 \end{Bmatrix}. \tag{A-4}$$

Then

$$\overline{p}(z_1, \ldots, z_n | z_{n+1} = \pm 1) = \int_{A(z_{n+1})}^{B(z_{n+1})} dx_{n+1} |x_{n+1}|^n \ p(|x_{n+1}| z_1, \ldots, |x_{n+1}| z_n, x_{n+1})$$

$$= \int_0^\infty dx \ x^n \ p(xz_1, \ldots, xz_n, \pm x) . \tag{A-5}$$

A similar relation holds for $\overline{q}$ in terms of q. Then (A-2) becomes

$$LR(z_1, \ldots, z_n | z_{n+1} = \pm 1) = \frac{\int_0^\infty dx \ x^n \ p(xz_1, \ldots, xz_n, \pm x)}{\int_0^\infty dx \ x^n \ q(xz_1, \ldots, xz_n, \pm x)} . \tag{A-6}$$

This is a general relation for the LR and indicates how the scaled samples $\{z_k\}$ should be processed.

Let us apply the general relation (A-6) to the following example, which is actually more general than the one considered in the main body of the text. Let

$$p(x_1, \ldots, x_n, x_{n+1}) = \prod_{k=1}^{n+1} \left\{ (2\pi\sigma^2)^{-1/2} \exp\left[- \frac{(x_k - m_k)^2}{2\sigma^2} \right] \right\}. \tag{A-7}$$

The PDF q is given by the same formula with $\{m_k\}$ set equal to zero. The $\{m_k\}$ need not be equal, but they are assumed known for the moment; σ^2 is unknown. Define

$$S = \sum_{k=1}^{n} z_k^2, \quad L = \sum_{k=1}^{n} m_k z_k, \quad U = \sum_{k=1}^{n+1} m_k^2,$$

$$\beta_\pm = \frac{L \pm m_{n+1}}{\sqrt{S+1}\,\sigma}. \tag{A-8}$$

Then the numerator of (A-6) is given by

$$\left(2\pi\sigma^2\right)^{-(n+1)/2} \int_0^\infty dx\, x^n \exp\left[-\frac{x^2(S+1) - 2x(L \pm m_{n+1}) + U}{2\sigma^2}\right]$$

$$= \left[2\pi(S+1)\right]^{-(n+1)/2} \exp\left[-\frac{U}{2\sigma^2}\right] \int_0^\infty dt\, t^n \exp\left[-\frac{1}{2}t^2 + \beta_\pm t\right]. \tag{A-9}$$

The denominator of (A-6) follows from (A-9) by setting $\{m_k\}$ equal to zero. Then (A-6) becomes

$$LR(z_1, \ldots, z_n | z_{n+1} = \pm 1) = \exp\left[-\frac{U}{2\sigma^2}\right] \frac{\int_0^\infty dt\, t^n \exp\left[-\frac{1}{2}t^2 + \beta_\pm t\right]}{\int_0^\infty dt\, t^n \exp\left[-\frac{1}{2}t^2\right]}. \tag{A-10}$$

Without having to evaluate the integrals, (A-10) is seen to be monotonically increasing with $\beta_\pm$, as depicted in Fig. A-1. Therefore, comparison of the LR with a threshold is equivalent to comparison of $\beta_\pm$ with a threshold. From (A-8), this is

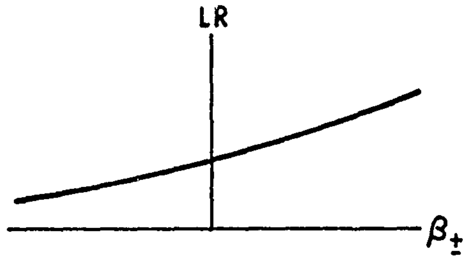

Fig. A-1. Dependence of the Likelihood Ratio on $\beta_\pm$

$$\frac{L \pm m_{n+1}}{\sqrt{S+1}} \gtrless \nu_1 , \qquad \text{(A-11)}$$

since σ, although unknown, is positive. Using (A-8) and (A-1) we express

$$L \pm m_{n+1} = \sum_{k=1}^{n} m_k \frac{x_k}{|x_{n+1}|} + m_{n+1} \frac{x_{n+1}}{|x_{n+1}|} = \frac{1}{|x_{n+1}|} \sum_{k=1}^{n+1} m_k x_k$$

and

$$(S + 1)^{1/2} = \left(\sum_{k=1}^{n} \frac{x_k^2}{x_{n+1}^2} + 1 \right)^{1/2} = \frac{1}{|x_{n+1}|} \left(\sum_{k=1}^{n+1} x_k^2 \right)^{1/2} . \qquad \text{(A-12)}$$

Therefore the LR test (A-11) becomes

$$\sum_{k=1}^{n+1} m_k x_k \gtrless \nu_1 \left(\sum_{k=1}^{n+1} x_k^2 \right)^{1/2} . \qquad \text{(A-13)}$$

This is the optimum invariant test if $\{m_k\}$ are known; $\{m_k\}$ can be positive, negative, or of mixed polarities. Threshold ν_1 is chosen for specified P_F.

The LR test in (A-13) can be expressed as

$$\sum_{k=1}^{n+1} \frac{m_k}{|m_{max}|} x_k \gtrless \nu_2 \left(\sum_{k=1}^{n+1} x_k^2 \right)^{1/2} , \tag{A-14}$$

which does not require knowledge of <u>absolute</u> signal level, but only relative levels. This is a UMP test under scaling invariance for unknown signal level provided the relative signal strengths $\{m_k/|m_{max}|\}_1^{n+1}$, which can be mixed positive and negative, are known.

A special case of (A-14) is afforded by

$$m_k = m > 0 \quad \text{for} \quad 1 \le k \le M; \quad \text{otherwise,} \quad m_k = 0 , \tag{A-15}$$

where m is unknown. Then there follows

$$\sum_{k=1}^{M} x_k \gtrless \nu_2 \left(\sum_{k=1}^{n+1} x_k^2 \right)^{1/2} . \tag{A-16}$$

There is no need to know m; this is a UMP test under scaling invariance and is recognized as the test shown in Table 1, with the identification of $n + 1 = M + N$.

It is worth noting that if only the n ratios

$$y_k = \frac{x_k}{x_{n+1}} , \quad 1 \le k \le n , \tag{A-17}$$

are available for making decisions, the LR is given by

$$\mathrm{LR}(y_1, \ldots, y_n) = \frac{\int_{-\infty}^{\infty} dx\, |x|^n\, p(xy_1, \ldots, xy_n, x)}{\int_{-\infty}^{\infty} dx\, |x|^n\, q(xy_1, \ldots, xy_n, x)}. \qquad \text{(A-18)}$$

From the example in (A-7) there follows

$$\mathrm{LR}(y_1, \ldots, y_n) = \exp\left[-\frac{U}{2\sigma^2}\right] \frac{\int_{-\infty}^{\infty} dt\, |t|^n \exp\left[-\frac{1}{2} t^2\right] \cosh(\beta_+ t)}{\int_{-\infty}^{\infty} dt\, |t|^n \exp\left[-\frac{1}{2} t^2\right]}, \qquad \text{(A-19)}$$

which is even and monotonically increasing in β_+. Comparison of the LR with a threshold is equivalent to comparison of $|\beta_+|$ with a threshold. By use of (A-8) and (A-12) the comparison becomes

$$\left| \sum_{k=1}^{n+1} m_k x_k \right| \gtrless \gamma \left(\sum_{k=1}^{n+1} x_k^2 \right)^{1/2}. \qquad \text{(A-20)}$$

Lack of knowledge about the polarity of the samples requires taking a magnitude of the linear sum. This test will perform more poorly than (A-13) because it is operating with less knowledge.

The procedure in this appendix is useful if it results in a UMP test with respect to the parameters in the PDFs. Although this happens for the example considered in this report, it is not generally the case, and realization of a specified P_F is not always attainable. Thus, it is often necessary to resort to a different approach, such as the principle of ML, that eliminates unknown parameters.

Appendix B

MATRIX MANIPULATIONS: CHARACTERISTIC VALUES, CHARACTERISTIC VECTORS, AND NORMALIZED MODAL MATRIX

Let $\mathbf{A}$ be an $M \times M$ Hermitian matrix of complex elements. Let*

$$\mathbf{A}\mathbf{e}_\ell = \lambda_\ell \mathbf{e}_\ell, \quad 1 \le \ell \le M, \tag{B-1}$$

where $\mathbf{e}_\ell$ is an $M \times 1$ column matrix, λ_ℓ is the ℓ-th characteristic value of $\mathbf{A}$, and $\mathbf{e}_\ell$ is the ℓ-th (right) characteristic vector of $\mathbf{A}$. Equation (B-1) can be written

$$\sum_{j=1}^{M} a_{ij} e_{\ell j} = \lambda_\ell e_{\ell i}, \quad 1 \le \ell, \ i \le M. \tag{B-2}$$

Let $\mathbf{Q} = [\mathbf{e}_1 \ldots \mathbf{e}_M]$; this is the normalized modal matrix of $\mathbf{A}$ [Ref. 9, pp. 37-39 and p. 47]. Also let

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_M \end{bmatrix}. \tag{B-3}$$

Then (A-1) can be expressed as the following sequence of equations:

$$\mathbf{A}\mathbf{e}_\ell = \lambda_\ell \mathbf{e}_\ell, \ 1 \le \ell \le M,$$

$$[\mathbf{A}\mathbf{e}_1 \cdots \mathbf{A}\mathbf{e}_M] = [\lambda_1 \mathbf{e}_1 \cdots \lambda_M \mathbf{e}_M],$$

$$\mathbf{A}[\mathbf{e}_1 \ldots \mathbf{e}_M] = [\mathbf{e}_1 \ldots \mathbf{e}_M]\,\boldsymbol{\lambda},$$

$$\mathbf{A}\mathbf{Q} = \mathbf{Q}\boldsymbol{\lambda},$$

$$\mathbf{A} = \mathbf{Q}\boldsymbol{\lambda}\mathbf{Q}^{-1}. \tag{B-4}$$

*See, for example, Ref. 9, secs. 1.11 and 1.16.

Now

$$\mathbf{Q}^H\mathbf{Q} = \begin{bmatrix} \mathbf{e}_1^H \\ \vdots \\ \mathbf{e}_M^H \end{bmatrix} [\mathbf{e}_1 \ldots \mathbf{e}_M] = \mathbf{I}\,, \tag{B-5}$$

since the characteristic vectors are orthonormal [Ref. 9, p. 44]. Therefore $\mathbf{Q}^H = \mathbf{Q}^{-1}$, and (B-4) can be expressed as

$$\mathbf{A} = \mathbf{Q}\boldsymbol{\lambda}\mathbf{Q}^H \ \ (\neq \mathbf{Q}^H\boldsymbol{\lambda}\mathbf{Q})\,. \tag{B-6}$$

Alternatively, we can express

$$\boldsymbol{\lambda} = \mathbf{Q}^H\mathbf{A}\mathbf{Q}\,. \tag{B-7}$$

The j, k-th term of $\mathbf{Q}^H\mathbf{Q}$ is

$$\mathbf{e}_j^H\mathbf{e}_k = \sum_{\ell=1}^{M} e_{j\ell}^* e_{k\ell} = \delta_{jk} = \sum_{\ell=1}^{M} e_{j\ell} e_{k\ell}^*\,, \tag{B-8}$$

and the j, k-th term of $\mathbf{Q}\mathbf{Q}^H$ is

$$\sum_{\ell=1}^{M} e_{\ell j} e_{\ell k}^* = \delta_{jk} = \sum_{\ell=1}^{M} e_{\ell j}^* e_{\ell k}\,. \tag{B-9}$$

From (B-4), the following sequence of equations holds:

$$\mathbf{Q}^H\mathbf{A} = \boldsymbol{\lambda}\mathbf{Q}^H\,,$$

$$\begin{bmatrix} \mathbf{e}_1^H \\ \vdots \\ \mathbf{e}_M^H \end{bmatrix} \mathbf{A} = \boldsymbol{\lambda} \begin{bmatrix} \mathbf{e}_1^H \\ \vdots \\ \mathbf{e}_M^H \end{bmatrix}, \quad \begin{bmatrix} \mathbf{e}_1^H\mathbf{A} \\ \vdots \\ \mathbf{e}_M^H\mathbf{A} \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{e}_1^H \\ \vdots \\ \lambda_M\mathbf{e}_M^H \end{bmatrix},$$

$$\mathbf{e}_\ell^H \mathbf{A} = \lambda_\ell \mathbf{e}_\ell^H, \quad 1 \le \ell \le M. \tag{B-10}$$

If we define the (left) characteristic vectors and numbers of $\mathbf{A}$ as $\mathbf{v}_\ell$ and γ_ℓ, respectively, then

$$\mathbf{v}_\ell^T \mathbf{A} = \gamma_\ell \mathbf{v}_\ell^T, \quad 1 \le \ell \le M, \tag{B-11}$$

where $\mathbf{v}_\ell$ is an $M \times 1$ column matrix. Comparison of (B-10) and (B-11) reveals that

$$\gamma_\ell = \lambda_\ell, \quad \mathbf{v}_\ell = \mathbf{e}_\ell^*, \quad 1 \le \ell \le M; \tag{B-12}$$

thus the characteristic numbers are equal, whereas the characteristic vectors are conjugates.

Appendix C

REDUCTION OF HERMITIAN FORM

Two problems will be addressed in this appendix: converting a Hermitian form to a sum of squares of uncorrelated RVs and determining the statistics of the Hermitian form when the RVs are Gaussian.

Let $\mathbf{X}$ be an $M \times 1$ column matrix of complex RVs, and let $\mathbf{B}$ be an $M \times M$ Hermitian matrix. We do not assume that $\mathbf{X}$ is a collection of Gaussian RVs, so that we allow complete generality in the first problem. The Hermitian form of interest is

$$F(\mathbf{X}) = \mathbf{X}^H \mathbf{B} \mathbf{X} \ . \tag{C-1}$$

Since $\mathbf{B}$ is Hermitian, the form F is real. Let

$$E\{\mathbf{X}\} \equiv \mathbf{m} \tag{C-2}$$

be the means of the RVs, and

$$\mathrm{Cov}\{\mathbf{X}\} = E\{(\mathbf{X}-\mathbf{m})(\mathbf{X}-\mathbf{m})^H\} \equiv \mathbf{K} = \mathbf{K}^H \tag{C-3}$$

be the covariance matrix of the RVs; $\mathbf{m}$ and $\mathbf{K}$ are arbitrary, except that $\mathbf{K}$ must be an allowable covariance matrix (that is, nonnegative definite).

We define transformed RVs $\mathbf{V}$ by

$$\mathbf{V} = \mathbf{Z}\mathbf{K}^{-1/2}\mathbf{X} \ , \tag{C-4}$$

where $\mathbf{Z}$ is $M \times M$ and is yet to be chosen. From (C-2) and (C-3),

$$E\{\mathbf{V}\} = \mathbf{Z}\mathbf{K}^{-1/2}\mathbf{m} \ ,$$

$$\mathrm{Cov}\{\mathbf{V}\} = E\left\{\mathbf{Z}\mathbf{K}^{-1/2}(\mathbf{X}-\mathbf{m})(\mathbf{X}-\mathbf{m})^H\mathbf{K}^{-1/2}\mathbf{Z}^H\right\} = \mathbf{Z}\mathbf{Z}^H \ . \tag{C-5}$$

Preceding page blank

This last matrix is desired to be diagonal, for then the RVs $\mathbf{V}$ will be uncorrelated. At the same time, the quantity

$$\mathbf{V}^H\mathbf{Y}\mathbf{V} = \mathbf{X}^H\mathbf{K}^{-1/2}\mathbf{Z}^H\mathbf{Y}\mathbf{Z}\mathbf{K}^{-1/2}\mathbf{X} \tag{C-6}$$

will be set equal to $F(\mathbf{X})$ by choice of $\mathbf{Y}$, which is also $M \times M$. In order to accomplish this, we must have

$$\mathbf{K}^{-1/2}\mathbf{Z}^H\mathbf{Y}\mathbf{Z}\mathbf{K}^{-1/2} = \mathbf{B}, \tag{C-7}$$

or

$$\mathbf{Z}^H\mathbf{Y}\mathbf{Z} = \mathbf{K}^{1/2}\mathbf{B}\mathbf{K}^{1/2} \equiv \mathbf{A} = \mathbf{A}^H. \tag{C-8}$$

But $\mathbf{A}$ is a known matrix, since $\mathbf{K}$ and $\mathbf{B}$ are known; solve for $\mathbf{A}$'s normalized modal matrix $\mathbf{Q}$ and characteristic value matrix $\boldsymbol{\lambda}$. Then, from (B-6),

$$\mathbf{A} = \mathbf{Q}\boldsymbol{\lambda}\mathbf{Q}^H. \tag{C-9}$$

If we choose

$$\mathbf{Z} = \mathbf{Q}^H, \quad \mathbf{Y} = \boldsymbol{\lambda}, \tag{C-10}$$

(C-8) is certainly satisified. Also, (C-5) becomes, by the use of (B-5),

$$\mathrm{Cov}\{\mathbf{V}\} = \mathbf{Q}^H\mathbf{Q} = \mathbf{I}, \tag{C-11}$$

which is diagonal as desired. At the same time, (C-6) yields

$$F(\mathbf{X}) = \mathbf{V}^H\mathbf{Y}\mathbf{V} = \mathbf{V}^H\boldsymbol{\lambda}\mathbf{V} = \sum_{\ell=1}^{M} \lambda_\ell \left|v_\ell\right|^2. \tag{C-12}$$

Thus F is a sum of (magnitude) squares of uncorrelated RVs $\mathbf{V}$. The only remaining statistics are the means

$$E\{\mathbf{V}\} = \mathbf{Q}^H \mathbf{K}^{-1/2} \mathbf{m} \equiv \boldsymbol{\mu}, \tag{C-13}$$

obtained by using (C-5) and (C-10).

To summarize, $\mathbf{A}$ is evaluated from $\mathbf{K}$ and $\mathbf{B}$ according to (C-8); $\mathbf{Q}$ and $\boldsymbol{\lambda}$ are computed (see Appendix B); then F is given by (C-12), where RVs $\mathbf{V}$ have the covariance matrix $\mathbf{I}$ and means $\boldsymbol{\mu}$ (C-13).

Now we specialize the above results to the case in which RVs $\mathbf{X}$ are Gaussian and real, and $\mathbf{B}$ is a real symmetric matrix. Since $\mathbf{V}$ is a linear transformation of $\mathbf{X}$ (see (C-4)), RVs $\mathbf{V}$ are Gaussian; they are also real because $\mathbf{A}$ is real. From (C-11) and (C-13), the PDF of $\mathbf{V}$ is

$$p(\mathbf{V}) = \prod_{\ell=1}^{M} \left\{ (2\pi)^{-1/2} \exp\left[-\frac{1}{2}(v_\ell - \mu_\ell)^2\right] \right\}. \tag{C-14}$$

The characteristic function (CF) of form F then follows, from (C-12), as

$$f(\xi) = E\{\exp(i\,\xi\, F(\mathbf{X}))\} = E\{\exp(i\,\xi \sum_{\ell=1}^{M} \lambda_\ell v_\ell^2)\}$$

$$= \prod_{\ell=1}^{M} \left\{ \int dv_\ell \, \exp(i\xi\lambda_\ell v_\ell^2)(2\pi)^{-1/2} \exp\left[-\frac{1}{2}(v_\ell - \mu_\ell)^2\right] \right\}$$

$$= \prod_{\ell=1}^{M} \left\{ (1 - i2\lambda_\ell \xi)^{-1/2} \exp\left(\frac{i\mu_\ell^2 \lambda_\ell \xi}{1 - i2\lambda_\ell \xi} \right) \right\}, \tag{C-15}$$

where $\boldsymbol{\lambda}$ are the characteristic numbers of $\mathbf{A}$ given by (C-8), $\boldsymbol{\mu}$ is given by (C-13), and $\mathbf{Q}$ is the normalized modal matrix of $\mathbf{A}$. This is the desired result.

The PDF of RV F can be obtained via a Fast Fourier Transform of the CF f, and the cumulative distribution of RV F can be found by means of the techniques described in Refs. 10 and 11.

As a special case of (C-15), consider

$$\lambda_\ell = \lambda, \quad 1 \le \ell \le M \ (\lambda > 0) . \tag{C-16}$$

The CF of the PDF

$$\frac{1}{2\lambda} \exp\left(-\frac{\mu^2}{2} - \frac{x}{2\lambda}\right)\left(\frac{x}{\mu^2\lambda}\right)^{\nu/2} I_\nu\left(\mu\sqrt{\frac{x}{\lambda}}\right), \quad x > 0 \quad (\lambda > 0) \tag{C-17}$$

is [Ref. 12, Eq. (3.433)]

$$\exp\left(\frac{i\mu^2\lambda\xi}{1 - i2\lambda\xi}\right)\left(1 - i2\lambda\xi\right)^{-\nu-1} . \tag{C-18}$$

Therefore the PDF of F follows, from (C-15), as

$$\frac{1}{2\lambda} \exp\left(-\frac{\mu_T^2}{2} - \frac{F}{2\lambda}\right)\left(\frac{F}{\mu_T^2\lambda}\right)^{M/4-1/2} I_{M/2-1}\left(\mu_T\sqrt{\frac{F}{\lambda}}\right), \quad F > 0 , \tag{C-19}$$

where

$$\mu_T^2 \equiv \sum_{\ell=1}^{M} \mu_\ell^2 . \tag{C-20}$$

Appendix D

DERIVATION OF DETECTION PROBABILITY

Define the RVs

$$L_x = \sum_{i=1}^{M} x_i, \quad S_x = \sum_{i=1}^{M} x_i^2, \quad S_y = \sum_{j=1}^{N} y_j^2 . \tag{D-1}$$

The GLR test (10) can then be expressed as:

$$\text{choose } H_1 \text{ if } L_x^2 - r^2 M(S_x + S_y) > 0 \text{ and } L_x > 0; \text{ choose } H_0 \text{ otherwise.} \tag{D-2}$$

Now define

$$q = L_x^2 - r^2 M(S_x + S_y) . \tag{D-3}$$

Then the probability of detection is

$$P_D = \text{Prob}\,(q > 0, \; L_x > 0 \,|\, H_1) . \tag{D-4}$$

In order to evaluate P_D, the joint PDF of q and L_x is required. Using (D-1) in (D-3) yields

$$q = \mathbf{X}^T \mathbf{B} \mathbf{X} - r^2 M \sum_{j=1}^{N} y_j^2 , \tag{D-5}$$

where

$$\mathbf{X}^T = [x_1 \ldots x_M], \quad B_{k\ell} = \begin{Bmatrix} 1 - r^2 M, & k = \ell \\ 1, & k \neq \ell \end{Bmatrix} . \tag{D-6}$$

To reduce the quadratic form in (D-5) to a sum of squares of uncorrelated RVs, we rely on the results and notation of Appendixes B and C to accomplish this task without duplication of effort. From (2) and (3),

$$E\{\mathbf{X}|H_k\} = m_k \mathbf{1}, \quad k = 0 \text{ or } 1, \tag{D-7}$$

where

$$m_k = \begin{Bmatrix} 0, & k = 0 \\ m, & k = 1 \end{Bmatrix}, \quad \mathbf{1}^T = [1 \ldots 1] . \tag{D-8}$$

The index k, which equals 0 or 1, indicates the two hypotheses H_0 and H_1, respectively. The covariance of $\mathbf{X}$ under hypothesis H_k is, from (2) and (3),

$$\mathbf{K}_k = \mathrm{Cov}\{\mathbf{X}|H_k\} = \sigma^2 \mathbf{I} \equiv \mathbf{K}, \quad k = 0 \text{ or } 1. \tag{D-9}$$

Thus, the covariance matrix is the same for both hypotheses. It then follows from (D-9) that the matrix of (C-8) is given by

$$\mathbf{K}_k^{1/2} \mathbf{B} \mathbf{K}_k^{1/2} = \sigma^2 \mathbf{B} \equiv \mathbf{A}, \quad k = 0 \text{ or } 1. \tag{D-10}$$

The characteristic numbers of $\mathbf{A}$ are [Ref. 13, Eqs. (58) and (60)]

$$\lambda_1 = \sigma^2 (1-r^2)M, \quad \lambda_2 = \ldots = \lambda_M = -\sigma^2 r^2 M . \tag{D-11}$$

Also, from (C-13), (D-7), (D-8), and (D-9), under hypothesis H_k,

$$\boldsymbol{\mu}_k = \frac{m_k}{\sigma} \mathbf{Q}^T \mathbf{1} . \tag{D-12}$$

Letting $\mu_{k\ell}$ be the ℓ-th component of $\boldsymbol{\mu}_k$,

$$\mu_{k\ell} = \frac{m_k}{\sigma} \mathbf{e}_\ell^T \mathbf{1} = \frac{m_k}{\sigma} \mathbf{1}^T \mathbf{e}_\ell . \tag{D-13}$$

In order to evaluate these quantities, we refer to (B-1) and (D-11) to obtain

$$\mathbf{A}\mathbf{e}_1 = \sigma^2(1-r^2)M\mathbf{e}_1 ,$$

$$\mathbf{A}\mathbf{e}_\ell = -\sigma^2 r^2 M\mathbf{e}_\ell , \quad 2 \le \ell \le M . \tag{D-14}$$

The solution for $\mathbf{e}_1$ is

$$\mathbf{e}_1 = M^{-1/2}\mathbf{1} , \tag{D-15}$$

from substitution in the top equation of (D-14); and using (D-10) and (D-6):

$$\mathbf{A}\mathbf{e}_1 = \mathbf{A}M^{-1/2}\mathbf{1} = M^{-1/2}\sigma^2\mathbf{B}\mathbf{1} = M^{-1/2}\sigma^2(1-r^2)M\mathbf{1} = \sigma^2(1-r^2)M\mathbf{e}_1 = \lambda_1\mathbf{e}_1 . \tag{D-16}$$

Therefore, from (D-13) and (D-8),

$$\mu_{k1} = \frac{m_k}{\sigma}\mathbf{1}^T\mathbf{e}_1 = \sqrt{M}\frac{m_k}{\sigma} = \sqrt{M}\frac{m}{\sigma}\,\delta_{k1} = d_T\delta_{k1} , \tag{D-17}$$

where we have defined

$$d_T = \sqrt{M}\,\frac{m}{\sigma} . \tag{D-18}$$

This quantity can be interpreted as a (voltage) SNR of the RV on the left side of the GLR test (see discussion of (16)).

In order to evaluate $\mu_{k\ell}$ for $\ell \ge 2$, we premultiply the second equation of (D-14) by $\mathbf{1}^T$:

$$\mathbf{1}^T\mathbf{A}\mathbf{e}_\ell = -\sigma^2 r^2 M\,\mathbf{1}^T\mathbf{e}_\ell , \quad 2 \le \ell \le M . \tag{D-19}$$

But

$$\mathbf{1}^T\mathbf{A} = (\mathbf{A}\mathbf{1})^T = \sigma^2(1-r^2)M\,\mathbf{1}^T \tag{D-20}$$

from (D-16), so (D-19) becomes

$$\sigma^2(1-r^2)M\mathbf{1}^T\mathbf{e}_\ell = -\sigma^2 r^2 M\mathbf{1}^T\mathbf{e}_\ell, \quad 2 \le \ell \le M,$$

$$\mathbf{1}^T\mathbf{e}_\ell = 0, \quad 2 \le \ell \le M. \tag{D-21}$$

Combining (D-13), (D-17), and (D-21),

$$\mu_{k\ell} = \begin{Bmatrix} d_T\delta_{k1}, & \ell = 1 \\ 0, & 2 \le \ell \le M \end{Bmatrix}, \quad k = 0 \text{ or } 1. \tag{D-22}$$

Therefore (C-13) becomes

$$E\{v_1\} = d_T\delta_{k1}, \quad E\{v_\ell\} = 0, \quad 2 \le \ell \le M. \tag{D-23}$$

The RV q in (D-5) is now expressible as

$$q = \sum_{\ell=1}^{M}\lambda_\ell v_\ell^2 - r^2MS_y = \sigma^2(1-r^2)Mv_1^2 - \sigma^2r^2M\sum_{\ell=2}^{M}v_\ell^2 - r^2M\sum_{j=1}^{N}y_j^2$$

$$= \sigma^2M\left[(1-r^2)v_1^2 - r^2\left(\sum_{\ell=2}^{M}v_\ell^2 + \sum_{j=1}^{N}y_j^2/\sigma^2\right)\right], \tag{D-24}$$

from (C-12), (D-11), and (D-1). But since

$$\mathbf{V} = \mathbf{Z}\mathbf{K}^{-1/2}\mathbf{X} = \frac{1}{\sigma}\mathbf{Q}^T\mathbf{X} \tag{D-25}$$

from (C-4), (C-10), and (D-9), it follows from (D-15) that

$$v_1 = \frac{1}{\sigma}\mathbf{e}_1^T\mathbf{X} = \frac{1}{\sigma}M^{-1/2}\mathbf{1}^T\mathbf{X} = \frac{1}{\sqrt{M}\,\sigma}\sum_{i=1}^{M}x_i. \tag{D-26}$$

Defining $z_j = y_j/\sigma$, (D-4) becomes, by using (D-1), (D-24), and (D-26),

$$P_D = \text{Prob}\left(v_1 > \frac{r}{\sqrt{1-r^2}}\left(\sum_{\ell=2}^{M} v_\ell^2 + \sum_{j=1}^{N} z_j^2\right)^{1/2} \middle| H_1\right). \tag{D-27}$$

But all the RVs **V** and **Y** are independent and Gaussian. Therefore, using (C-11), (D-23), and (3) gives

$$p_1(v_1) = (2\pi)^{-1/2} \exp(-(v_1 - d_T)^2/2) ,$$

$$p_1(v_\ell) = (2\pi)^{-1/2} \exp(-v_\ell^2/2), \quad 2 \le \ell \le M ,$$

$$p_1(z_j) = (2\pi)^{-1/2} \exp(-z_j^2/2), \quad 1 \le j \le N . \tag{D-28}$$

It follows that the PDF of

$$w \equiv \left(\sum_{\ell=2}^{M} v_\ell^2 + \sum_{j=1}^{N} z_j^2\right)^{1/2} \tag{D-29}$$

is [Ref. 14, Eq. (26.4)]

$$p(w) = \frac{w^K \exp(-w^2/2)}{2^{(K-1)/2}\,\Gamma\left(\frac{K+1}{2}\right)}, \quad w > 0 , \tag{D-30}$$

where

$$K \equiv M + N - 2 . \tag{D-31}$$

Then (D-27) becomes

$$P_D = \text{Prob}\left(v_1 > \frac{r}{\sqrt{1-r^2}}w \middle| H_1\right)$$

$$= \int_0^\infty dw\, p(w) \int_{\frac{r}{\sqrt{1-r^2}}w}^\infty dv_1\, p_1(v_1)$$

$$= \int_0^\infty dw \frac{w^K \exp(-w^2/2)}{2^{(K-1)/2}\,\Gamma\left(\frac{K+1}{2}\right)}\, \Phi\left(d_T - \frac{r}{\sqrt{1-r^2}}w\right)$$

$$\equiv f(d_T, r, K)\,, \tag{D-32}$$

where

$$\Phi(x) \equiv \int_{-\infty}^x dt\, (2\pi)^{-1/2} \exp(-t^2/2) \equiv \int_{-\infty}^x dt\, \phi(t)\,. \tag{D-33}$$

The false-alarm probability follows by setting $d_T = 0$ (m=0) in (D-32).

The above derivation has tacitly assumed $M \geq 2$. However, it may readily be shown by a separate derivation that (D-32) is also correct for $M=1$, $N=1$. Therefore (D-32) holds for $K \geq 0$, as mentioned in the footnote to (14) in the main text.

Appendix E

REDUCTION OF EQUATION (D-32)

Interest here lies in evaluating (D-32):

$$P_D = f(d_T, r, K) = \int_0^\infty dw \frac{w^K \exp(-w^2/2)}{2^{(K-1)/2}\,\Gamma\left(\frac{K+1}{2}\right)} \Phi\left(d_T - \frac{r}{\sqrt{1-r^2}} w\right), \quad K \geq 0 . \qquad \text{(E-1)}$$

First, for $K = 0$,

$$f(d_T, r, 0) = \int_0^\infty dw \left(\frac{2}{\pi}\right)^{1/2} \exp(-w^2/2)\ \Phi\left(d_T - \frac{r}{\sqrt{1-r^2}} w\right)$$

$$= 2L(0, -\sqrt{1-r^2}\, d_T, -r) \qquad \text{(E-2)}$$

[Ref. 14, Eq. (26.3.3)]. Since this last function, L, is fundamental [Ref. 14, Eq. (26.3.20)], (E-2) can not be reduced, and must be numerically evaluated. This problem is discussed in Appendix F.

For $K = 1$, integrate by parts, letting

$$u = \Phi\left(d_T - \frac{r}{\sqrt{1-r^2}} w\right), \quad dv = dw\ w \exp(-w^2/2), \qquad \text{(E-3)}$$

to obtain

$$f(d_T, r, 1) = \Phi(d_T) - r \exp(-(1-r^2)\, d_T^2/2)\ \Phi(r d_T) . \qquad \text{(E-4)}$$

For $K \geq 2$, integrate by parts, letting

$$u = w^{K-1}\, \Phi\left(d_T - \frac{r}{\sqrt{1-r^2}} w\right), \quad dv = dw\, w \exp(-w^2/2)\,, \tag{E-5}$$

to obtain

$$f(d_T, r, K) = f(d_T, r, K-2) - g(d_T, r, K-1), \quad K \geq 2\,, \tag{E-6}$$

where

$$g(d_T, r, K) \equiv \left[2^{K/2}\Gamma\left(\frac{K}{2}+1\right)\right]^{-1} \frac{r}{\sqrt{1-r^2}} \int_0^\infty dw\, w^K \exp(-w^2/2)\, \phi\left(d_T - \frac{r}{\sqrt{1-r^2}} w\right), \quad K \geq 0\,, \tag{E-7}$$

and ϕ is defined in (D-33). Using (E-4), (E-6) can be expressed as

$$f(d_T, r, K) = \left\{ \begin{array}{ll} \Phi(d_T) - \displaystyle\sum_{\substack{n=0 \\ n \text{ even}}}^{K-1} g(d_T, r, n), & K = 1, 3, 5, \ldots \\ f(d_T, r, 0) - \displaystyle\sum_{\substack{n=1 \\ n \text{ odd}}}^{K-1} g(d_T, r, n), & K = 2, 4, 6, \ldots \end{array} \right\}, \tag{E-8}$$

where we have noted that the last term in (E-4) is $g(d_T, r, 0)$; that is, by direct integration of (E-7) for $K = 0$,

$$g(d_T, r, 0) = r \exp(-(1-r^2)d_T^2/2)\, \Phi(r d_T)\,. \tag{E-9}$$

Now we have the problem of evaluating the g-functions in (E-7). Again, integrate by parts for $K \geq 1$, letting

$$u = w^{K-1}\,\phi\left(d_T - \frac{r}{\sqrt{1-r^2}}w\right), \quad dv = dw\, w\, \exp(-w^2/2)\,. \tag{E-10}$$

Then

$$\Big[uv\Big]_0^\infty = \left[-\exp(-w^2/2)\, w^{K-1}\,\phi\left(d_T - \frac{r}{\sqrt{1-r^2}}w\right)\right]_0^\infty = \phi(d_T)\,\delta_{k1}\,, \tag{E-11}$$

and there follows

$$g(d_T, r, K) = r\sqrt{1-r^2}\left[\frac{1}{\pi}\exp(-d_T^2/2)\delta_{k1} + \frac{d_T}{\sqrt{2}}\frac{\Gamma\left(\frac{K+1}{2}\right)}{\Gamma\left(\frac{K}{2}+1\right)}g(d_T, r, K-1)\right]$$

$$+\frac{K-1}{K}(1-r^2)\, g(d_T, r, K-2)\,, \tag{E-12}$$

where the last term is absent if $K = 1$. Therefore

$$g(d_T, r, 1) = r\sqrt{1-r^2}\left[\frac{1}{\pi}\exp(-d_T^2/2) + \sqrt{\frac{2}{\pi}}\,d_T\, g(d_T, r, 0)\right], \tag{E-13}$$

and

$$g(d_T, r, K) = r\sqrt{1-r^2}\,\frac{d_T}{\sqrt{2}}\frac{\Gamma\left(\frac{K+1}{2}\right)}{\Gamma\left(\frac{K}{2}+1\right)}\, g(d_T, r, K-1)$$

$$+\frac{K-1}{K}(1-r^2)\, g(d_T, r, K-2), \quad K \geq 2\,. \tag{E-14}$$

Thus, $g(d_T, r, 0)$ is evaluated from (E-9), $g(d_T, r, 1)$ from (E-13), and, for $K \geq 2$, $g(d_T, r, K)$ from (E-14). Then $P_D = f(d_T, r, K)$ is evaluated from (E-2) and (E-8). For K odd, nothing worse than a Φ-function need be evaluated.*

---

*See, for example, Ref. 14, Eq. (26.2.17).

Also, the recurrence, (E-14), contains only positive terms and thereby retains accuracy for large K.

The false alarm probability is obtained by setting $d_T = 0$ in the above results. From (E-1),

$$P_F = f(0, r, K) . \tag{E-15}$$

Now, for $K = 0$, using (E-2) and Ref. 14, Eq. (26.3.19), gives

$$f(0, r, 0) = 2L(0, 0, -r) = \frac{1}{2} - \frac{1}{\pi} \text{arc}\sin(r) . \tag{E-16}$$

The expressions for $g(0, r, K)$ take on very simple forms, as may be seen from (E-9), (E-13), and (E-14). Then

$$f(0, r, K) = \begin{cases} \dfrac{1}{2} - \dfrac{1}{2} r \displaystyle\sum_{\ell=0}^{(K-1)/2} \frac{(2\ell)!}{(\ell!)^2} \left[\frac{1 - r^2}{4}\right]^\ell , & K = 1, 3, 5, \ldots \\ \\ \dfrac{1}{2} - \dfrac{1}{\pi} \text{arc}\sin(r) - \dfrac{1}{\pi} r \sqrt{1-r^2} \displaystyle\sum_{\ell=0}^{K/2-1} \frac{(\ell!)^2}{(2\ell+1)!} [4(1-r^2)]^\ell , & K = 2, 4, 6, \ldots \end{cases} \tag{E-17}$$

Equations (E-16) and (E-17) constitute closed-form solutions for the false-alarm probability (E-15). For each value of K, r can be selected to realize a prescribed P_F. Then P_D can be evaluated versus d_T from (E-2) and (E-8).

Appendix F

EVALUATION OF EQUATION (E-2)

We have, from (E-2),

$$f(d_T, r, 0) = \int_0^\infty dw \left(\frac{2}{\pi}\right)^{1/2} \exp(-w^2/2) \int_{\frac{rw}{\sqrt{1-r^2}}}^{\infty} dv (2\pi)^{-1/2} \exp(-\tfrac{1}{2}(v-d_T)^2) \,. \quad \text{(F-1)}$$

Let $w = \rho \cos\theta$, $v = \rho \sin\theta$. Then, expanding $\exp(d_T \rho \sin\theta)$ in a power series to eliminate cross-product terms,

$$f(d_T, r, 0) = \frac{1}{\pi} \exp(-d_T^2/2) \int_0^\infty d\rho\, \rho \int_{\arcsin(r)}^{\pi/2} d\theta\, \exp(-\tfrac{1}{2}\rho^2 + d_T \rho \sin\theta)$$

$$= \sum_{n=0}^{\infty} \left[\frac{1}{\pi} \exp(-d_T^2/2) \frac{d_T^n}{n!} \int_0^\infty d\rho\, \rho^{n+1} \exp(-\rho^2/2) \right] \left[\int_{\arcsin(r)}^{\pi/2} d\theta\, \sin^n\theta \right]$$

$$\equiv \sum_{n=0}^{\infty} A_n(d_T)\, B_n(r) \,. \quad \text{(F-2)}$$

In the ρ integration, integrate by parts using $u = \rho^n$, $dv = d\rho\, \rho \exp(-\rho^2/2)$; do likewise in the θ integration, using $u = \sin^{n-1}\theta$, $dv = d\theta \sin\theta$. The two following resulting recurrence relations enable rapid determination of the series (F-2):

$$A_o(d_T) = \frac{1}{\pi} \exp(-d_T^2/2), \quad A_1(d_T) = d_T\left(\frac{\pi}{2}\right)^{1/2} A_o(d_T),$$

$$A_n(d_T) = \frac{d_T^2}{n-1} A_{n-2}(d_T), \quad n \geq 2, \tag{F-3}$$

and

$$B_o(r) = \frac{\pi}{2} - \arcsin(r), \; B_1(r) = (1-r^2)^{1/2},$$

$$B_n(r) = \frac{r^{n-1}\left(1-r^2\right)^{1/2} + (n-1)B_{n-2}(r)}{n}, \quad n \geq 2. \tag{F-4}$$

REFERENCES

1. C. W. Helstrom, Statistical Theory of Signal Detection, second edition, Pergamon Press, N.Y., 1968.

2. L. L. Scharf and D. W. Lytle, "Signal Detection in Gaussian Noise of Unknown Level: An Invariance Application," Institute of Electrical and Electronics Engineers Transactions on Information Theory, vol. IT-17, no. 4, July 1971, pp. 404-411.

3. J. W. Carlyle and J. B. Thomas, "On Nonparametric Signal Detectors," Institute of Electrical and Electronics Engineers Transactions on Information Theory, vol. IT-10, no. 2, April 1964, pp. 146-152.

4. J. Capon, "A Nonparametric Technique for the Detection of a Constant Signal in Additive Noise," Institute of Radio Engineers WESCON Convention Record, vol. 3, part 4, 1959, pp. 92-103.

5. J. C. Hancock and D. G. Lainiotis, "On Learning and Distribution-Free Coincidence Detection Procedures," Institute of Electrical and Electronics Engineers Transactions on Information Theory, vol. IT-11, no. 2, April 1965, pp. 272-281.

6. R. L. Spooner, "On the Detection of a Known Signal in a Non-Gaussian Noise Process," Journal of the Acoustical Society of America, vol. 44, no. 1, January 1968, pp. 141-147.

7. H. L. Van Trees, Detection, Estimation, and Modulation Theory, Part 1, J. Wiley and Sons, N.Y., 1968.

8. E. L. Lehmann, Testing Statistical Hypotheses, J. Wiley and Sons, N.Y., 1959.

9. F. B. Hildebrand, Methods of Applied Mathematics, Prentice-Hall, N.Y., 1954.

10. A. H. Nuttall, Numerical Evaluation of Cumulative Probability Distribution Functions Directly from Characteristic Functions, NUSL Technical Report No. 1032, 11 August 1969; also Proceedings of the Institute of Electrical and Electronics Engineers, vol. 57, no. 11, November 1969, pp. 2071-2072.

11. A. H. Nuttall, Alternate Forms and Computational Considerations for Numerical Evaluation of Cumulative Probability Distributions Directly from Characteristic Functions, NUSC Technical Report No. NL-3012, 12 August 1970; also Proceedings of the Institute of Electrical and Electronics Engineers, vol. 58, no. 11, November 1970, pp. 1872-1873.

12. A. D. Wheelon, Tables of Summable Series and Integrals Involving Bessel Functions, Holden-Day, San Francisco, 1968.

13. A. H. Nuttall, "Minimum RMS Bandwidth of M Time-Limited Signals with Specified Code or Correlation Matrix," Institute of Electrical and Electronics Engineers Transactions on Information Theory, vol. IT-14, no. 5, September 1968, pp. 699-707.

14. Handbook of Mathematical Functions, U. S. Department of Commerce, National Bureau of Standards, Applied Mathematics Series No. 55, U. S. Government Printing Office, Washington, D.C., June 1964.